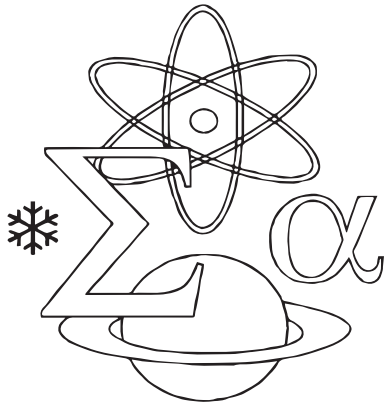
AJNTS

ALBANIAN JOURNAL OF NATURAL AND TECHNICAL SCIENCES



2025 (1)

XXX (62)

PUBLISHED BY ACADEMY OF SCIENCES OF ALBANIA

AJNTS First published in 1996

Editorial Board

Editor-in-Chief: Acad. **Neki Frashëri**, Academy of Sciences of Albania

Dhimitër Haxhimihali, Academy of Sciences of Albania
Floran Vila, Academy of Sciences of Albania
Ilia Mikerezi, Academy of Sciences of Albania
Alban Ylli, Institute of Public Health, Albania
Aleko Miho, University of Tirana, Albania
Andrea Maliqari, Polytechnic University of Tirana, Albania
Fatmir Hoxha, University of Tirana, Albania
Ylber Muceku, Institute of Geosciences, Polytechnic University of Tirana, Albania
Arben Merkoçi, Catalan Institute of Nanoscience and Nanotechnology, Spain
Arjan Durrësi, Purdue University, USA
Fioralba Cakoni, Rutgers University, USA
Mira Mezini, Technical University of Darmstadt, Germany
Arben Myrta, Certis Europe B.V, Utrecht, The Netherlands
Fetah Podvorica, Academy of Sciences and Arts of Kosovo
Andrea Pieroni, University of Gastronomic Sciences, Italy
Domenico Schiavone, University of Bari, Italy
Doncho Donev, University Ss. Cyril and Methodius, Republic of North Macedonia
Felix Unger, European Academy of Sciences and Arts, and Alma Mater Europaea
Giuseppe Baldassarre, University of Bari, Italy, Academy of Puglia
Massimo Chiappini, National Institute of Geophysics and Volcanology, Italy
Peter MacGrath, InterAcademy Partnership (IAP), Italy
Vlado Matevski, Macedonian Academy of Sciences and Arts

Science Editor
Msc. Blerina Shkreta
Academy of Sciences, Tirana, Albania
E-mail: akadshkreta@gmail.com

Aims and Scope
This Journal is a multidisciplinary publication devoted to all field of Natural and Technical Sciences. The Editor of AJNTS invites original contributions which should comprise previously unpublished results, data and interpretations.
Types of contributions to be published are: (1) research papers; (2) shorts communications; (3) reviews; (4) discussions; (5) book reviews;(6) announcements

ISSN 2074-0867

Graphic by: Dorina Muka, Printed by "Akademia"

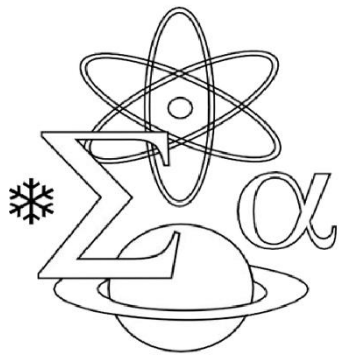
© Academy of Sciences of Albania, 2025 This journal operates under an Open Access model. Articles are freely accessible to the public immediately upon publication. The content is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), allowing users to share and adapt the work with proper attribution.

Copyright remains with the author(s), and no permission is required for non-commercial use, provided the original source is cited.

PUBLISHED BY THE ACADEMY OF SCIENCES OF ALBANIA

AJNTS

ALBANIAN JOURNAL OF NATURAL AND TECHNICAL SCIENCES



2025 (1)
XXX (62)

MODELING GEOMAGNETIC FIELD REVERSALS WITH COUPLED PARTICLES IN A DOUBLE-WELL POTENTIAL

Klaudio PEQINI
Department of Physics, Faculty of Natural Sciences, University of Tirana, Albania
Author for correspondence: klaudio.peqini@fshn.edu.al
KP, 0000-0002-8516-7381
https://doi.org/10.67537/ajnts3001001

ABSTRACT

We study a minimal model of geomagnetic field reversals, inspired by Molina-Cardin *et al*. (2021), in which two particles evolve in a double-well potential and interact through elastic coupling. One particle evolves slowly, while the other evolves more rapidly, representing distinct dynamical scales of the geodynamo. We explore two formulations of this system: a first-order diffusive (Brownian) version and a second-order inertial (Langevin) version. These formulations exhibit markedly different statistical behaviors: the Brownian model tends to generate unrealistically frequent reversals, especially when calibrated with limited datasets, whereas the Langevin model produces reversal sequences that align more closely with the long-term geomagnetic record. Using two different palaeomagnetic records and comparing the resulting reversal series with several extended polarity timescales, we demonstrate that the model's behavior is highly sensitive to the choice of palaeomagnetic record. While the diffusive version produces an unrealistic number of reversals, the Langevin version shows a qualitative consistency with polarity records. These findings highlight the need for high-quality, long-duration datasets and support the inclusion of inertial dynamics to suppress spurious reversal events. Although the two-particle framework provides insights into the interplay between fast and slow processes in the outer core, its simplicity imposes limitations. We conclude that future developments should aim to extend the model with additional degrees of freedom—potentially through coupling with spatially structured systems—to better capture the complexity of geomagnetic reversals observed in both palaeomagnetic data and numerical geodynamo simulations.

***Keywords*:** dynamo theory, particles and well, palaeomagnetic measurements, asymmetry of reversals, Brownian model

1. INTRODUCTION

The Earth's magnetic field (also known as the geomagnetic field) is one of the most extensively studied magnetic properties of astrophysical objects (Merrill *et al*. 1996). A wealth of direct measurements and historical data exists, covering a wide range of temporal and spatial scales—from thousands to millions of years (Valet and Meynadier, 1993; Guyodo and Valet, 1999; Valet *et al*. 2006; Guyodo and Valet, 2006; Ziegler and Constable, 2011). The field is generated by the geodynamo mechanism, which involves convective flows in the liquid outer core driven by thermal and compositional gradients (Backus *et al*. 1996; Kageyama and Sato, 1997). These flows interact in complex ways with the ambient magnetic field (Dormy and Soward, 2007; Duka *et al*. 2015), giving rise to a highly intricate system. Despite these complexities, the geomagnetic field is predominantly dipolar in nature (Merrill *et al*. 1996; Backus *et al*. 1996; Duka *et al*. 2015; Peqini *et al*. 2015), resembling the field of a giant bar magnet tilted approximately 11 degrees from the Earth's rotational axis. However, unlike a simple bar magnet, this field is not stable: it has reversed polarity numerous times throughout geological history.

Over geological timescales, polarity reversals—during which the magnetic North and South poles switch places—occur irregularly. The interval between successive reversals ranges from hundreds of thousands to several million years, representing some of the longest-timescale variations observed in the geomagnetic record (Backus *et al*. 1996). On average, reversals occur every 200,000 to 300,000 years, although the frequency has varied significantly over time (Valet and Meynadier, 1993; Backus *et al*. 1996; Guyodo and Valet, 2006; Olson *et al*. 2010; Mori *et al*. 2013; Duka *et al*. 2015; Peqini *et al*. 2015; Valet and Fournier, 2016). These fluctuations in reversal rates are striking (see Fig. 1), with intervals of frequent reversals alternating with superchrons—extended periods, such as the Cretaceous Normal Superchron, which lasted up to 38 million years with virtually no reversals (Cande and Kent, 1992).

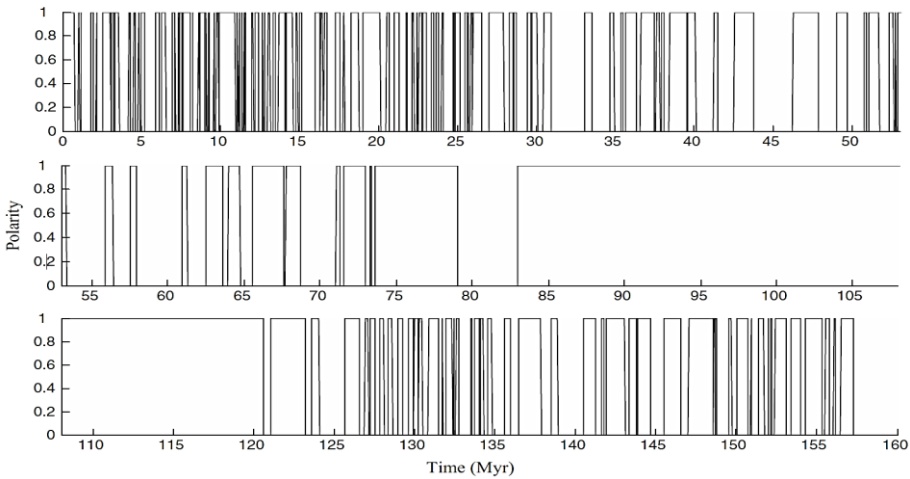


Fig. 1: The time series of the polarity reversals for the period of 157.5 Myr provided by Cande and Kent (1992). There it is shown the polarity of the Dipolar field for the last 157.5 million years. The time "0" refers to the present (adapted from Duka *et al.* 2015).

In recent years, low-dimensional geomagnetic models—simplified representations of the geodynamo—have received increasing attention in the scientific literature (Mori *et al.* 2013; Duka *et al.* 2015; Peqini *et al.* 2015; Molina-Cardin *et al.* 2021; Peqini, 2024). Rather than attempting to reproduce the full complexity of the geodynamo, these models focus on specific aspects of geomagnetic behavior, such as dipolar field reversals. The aim is to extract meaningful insights from an inherently complex system governed by intricate dynamical equations. Despite their simplicity, these models have proven successful in providing coherent explanations for key features of geomagnetic reversals. For example, they offer insight into the mechanisms underlying polarity switches (Mori *et al.* 2013; Duka *et al.* 2015) and explain the intermittent nature of reversals, often attributed to the intrinsic stochasticity of the system (Duka *et al.* 2015; Peqini *et al.* 2015; Molina-Cardin *et al.* 2021).

One low-dimensional model has successfully reproduced a key feature of the geomagnetic reversal mechanism: the slow decay of the dipolar field followed by its rapid recovery with opposite polarity (Molina-Cardin *et al.* 2021). This asymmetry is well documented in palaeomagnetic time series (Valet and Meynadier, 1993; Guyodo and Valet, 1999; Valet *et al.* 2005; Guyodo and Valet, 2006; Ziegler and Constable, 2011). In this

paper, we examine two existing versions of this model and assess whether they can account for additional observed patterns of geomagnetic reversals. Although low-dimensional models inherently possess limited scope, they nevertheless provide opportunities for further development to more accurately capture the underlying system. In this work, we explore such extensions and evaluate their potential to enhance the model's explanatory power.

The paper is organized as follows: in Section 2, we describe the model proposed by Molina-Cardín *et al.* (2021). Section 3 presents key results, while Section 4 discusses additional aspects of the model, including its strengths and limitations. This section also introduces several proposed modifications and extensions aimed at improving the model's performance. Finally, Section 5 offers concluding remarks and outlines potential directions for future research.

2. Particles in a well model

The *particles-in-a-well* model represents an interesting and relatively simple low-dimensional framework (see Fig. 1 in Molina-Cardín *et al.* 2021). Two versions of this model can be constructed: the diffusive version, as introduced by Molina-Cardín *et al*. (2021), and a second-order version derived from Langevin-type dynamical equations (Peqini, 2024).

An earlier and relatively successful model was that of a single particle in a double-well potential (Schmitt *et al*., 2001). In this model, the particle undergoes Brownian motion within a quartic potential of the form:

$$V(x) = ax^4 + bx^2 + c$$

Here the particle's position represents the magnitude of the dipolar geomagnetic moment, while its Brownian motion reflects the random-like temporal variations observed across multiple timescales (Merrill *et al*. 1996; Schmitt *et al*. 2001; Aubert *et al*. 2013; Buffett *et al*. 2013; Morzfeld and Buffett, 2019). Each minimum of the potential—i.e., each well—corresponds to a stable configuration of the dipolar magnetic field: the positive well represents the normal polarity state, and the negative well represents the reversed polarity state. Within this framework, geomagnetic reversals are interpreted as intermittent transitions of the particle between wells. However, although this model provides a compelling statistical interpretation of reversals, the inherent symmetry of the potential prevents

it from capturing the observed asymmetry between the slow decay and rapid recovery phases of real geomagnetic reversals.

Before presenting the mathematical formulation of the model, it is helpful to clarify the motivation underlying the particles-in-a-well framework. The two-particle approach adopted here extends a long-standing tradition of reducing the complex geodynamo dynamics into low-dimensional systems that capture essential features of geomagnetic polarity reversals. Gubbins (1988), for instance, suggested that reversals may arise from interactions between two competing dynamo modes: a dominant axial dipole and a secondary non-axisymmetric mode that can destabilize the system under certain conditions. This dual-mechanism perspective provides the conceptual foundation for our two-particle representation. In this analogy, one particle corresponds to the slow evolution of the dominant dipolar component, while the other represents faster, potentially destabilizing fluctuations associated with higher-order modes or turbulence. The elastic coupling between the two particles introduces dynamic feedback: perturbations in the fast mode can affect the stability of the slow mode, and vice versa. This interplay is essential for reversal dynamics, as it provides a mechanism through which internal fluctuations can trigger large-scale polarity switches.

Further theoretical support for this analogy comes from the distinction between geomagnetic excursions and full reversals, as articulated by Gubbins (1999). In that framework, excursions are interpreted as failed reversals—events potentially triggered by the same underlying mechanism but lacking sufficient perturbation amplitude or duration to fully destabilize the dipolar field. Our model is consistent with this view: depending on the strength and timing of stochastic forcing, the fast particle may either succeed or fail in driving the system across the potential barrier, leading to full reversals or limited excursions. This formulation enables a unified interpretation of both phenomena within a single stochastic system, where the transition probability is governed by the interplay of noise intensity, potential well asymmetry, and coupling strength.

Hoyng *et al*. (2002) further advanced the dual-mode approach by analyzing the statistical variability of the geomagnetic dipole in terms of nonlinear interactions between a stable dipolar mode and a more erratic secondary mode. Their results support the view that reversals are not purely random events but emerge from the interaction of modes with different symmetries and timescales. Our model reflects this structure by assigning distinct dynamical roles to each particle: one with slow,

overdamped dynamics representing the stable dipole, and the other with faster, potentially inertial behavior that introduces irregular forcing. This separation of timescales and roles reinforces the idea that reversals are emergent phenomena driven by internal dynamical coupling rather than externally triggered or purely stochastic processes. By incorporating both diffusive and Langevin formulations, our study further explores how different physical assumptions about the fast mode shape the statistical properties of the resulting reversal sequences.

In this contribution, we restrict our focus to the model's ability to qualitatively reproduce the reversal features reported in the literature. A subsequent paper will provide a more detailed quantitative analysis, with particular emphasis on measures used to characterize reversal asymmetry.

An elegant solution to the asymmetry problem was proposed by Molina-Cardín *et al*. (2021), who extended the earlier single-particle framework by introducing a second particle into the double-well potential, coupled to the first via a spring. Both particles undergo random motion but with different kinetic energies—effectively, distinct "temperatures." The fast particle represents short-term, high-frequency geomagnetic variability, such as secular variation and smaller-scale fluctuations (Guyodo and Valet, 1999; Aubert *et al*. 2013; Morzfeld and Buffett, 2019). By contrast, the slow particle captures long-term variations spanning geological timescales (Backus *et al*. 1996; Merrill *et al*. 1996; Valet *et al*. 2006; Ziegler and Constable, 2011; Wicht and Meduri, 2016). These two classes of timescales are not independent; rather, long-term behavior can be interpreted as a natural consequence of short-term dynamics (Backus *et al*. 1996; Duka *et al*. 2015). The coupling between particles is modeled through an elastic interaction, represented by a spring potential—hence the name of the model. Graphical illustrations and further technical details are provided in Molina-Cardín *et al*. (2021).

The physical and observational motivations discussed above lead to the following Lagrangian:

$$L = \frac{1}{2}\left[\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2\right] - \frac{1}{2}K\left(x_2 - x_1\right)^2 - V\left(x_1, x_2\right) \tag{1}$$

Here, the kinetic energy terms correspond to the two particles, each assumed to have unit mass. The elastic term models their interaction through a spring with stiffness constant *K*, while the potential energy term

$V(x_1, x_2)$ defines the double-well structure experienced by each particle. The gradient of this potential determines the force acting on the particles within the well. The system's dynamics are described by a set of Langevin-type equations, expressed as follows:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\left(dx_i/dt\right)}\right)=\frac{\partial L}{\partial x_i}-\gamma_i\frac{dx_i}{dt}+\sqrt{2k_BT_i\gamma_i}\chi_i,\quad i=1,2 \tag{2}$$

where the γ coefficients represent the drag experienced by each particle, and T_i denotes the temperature associated with each particle. In this context, a higher temperature corresponds to a faster particle. The random variable χ is drawn from the normal distribution with zero mean and unit variance. Following the same steps as in the previous section leads to the Langevin-type version:

$$\begin{cases}\dfrac{d^2x_1}{dt^2}=-\left.\dfrac{dV}{dx}\right|_{x_1}-\gamma_1\dfrac{dx_1}{dt}+K\left(x_2-x_1\right)+\sqrt{2k_BT_1\gamma_1}\chi_1\left(t\right)\\ \dfrac{d^2x_2}{dt^2}=-\left.\dfrac{dV}{dx}\right|_{x_2}-\gamma_2\dfrac{dx_2}{dt}+K\left(x_1-x_2\right)+\sqrt{2k_BT_2\gamma_2}\chi_2\left(t\right)\end{cases} \tag{3}$$

The diffusive model, obtained by neglecting the second time derivative, is given by:

$$\begin{cases}\dfrac{dx_1}{dt}=\dfrac{1}{\gamma_1}\left[-\left.\dfrac{dV}{dx}\right|_{x_1}+K\left(x_2-x_1\right)+\sqrt{2k_BT_1\gamma_1}\chi_1\left(t\right)\right]\\ \dfrac{dx_1}{dt}=\dfrac{1}{\gamma_2}\left[-\left.\dfrac{dV}{dx}\right|_{x_2}+K\left(x_1-x_2\right)+\sqrt{2k_BT_2\gamma_2}\chi_2\left(t\right)\right]\end{cases} \tag{4}$$

In both cases of V', indicates the partial derivative of the potential with respect to the corresponding variable. The variables x_1 and x_2 represent the contributions of the slow and fast particle, respectively, to the total dipolar moment X, also referred to here as the *axial component*. The output of the model is the axial component, $X(t)$, is calculated as below:

$$X(t) = wx_1(t) + (1-w)x_2(t) \quad (5)$$

The weight (w) factor indicates which particle contributes more to the output. For a consistent model, *w* must be less than 0.5, this ensuring that the slow particle predominantly determines the long-term the trend.

The potential function *V*(x), which represents the double-well structure, is obtained by fitting a quartic equation in *x* to the probability density function (PDF) derived from palaeomagnetic dipolar field datasets. These datasets record the geomagnetic dipolar field across geological periods ranging from thousands to millions of years (Valet and Meynadier, 1993; Guyodo and Valet, 1999; Valet *et al*. 2005; Guyodo and Valet, 2006; Ziegler and Constable, 2011). In the study by Molina-Cardín *et al*. (2021), palaeomagnetic records covering the last 800,000 years were used (Guyodo and Valet, 1999).

3. RESULTS

A typical realization of system (3) is shown in Fig. 2. The numerical implementation employs the fourth-order Runge–Kutta method, as discussed in Peqini (2024), which is equally applicable in this context. The fast particle, being at a higher temperature, experiences stronger fluctuations and frequently transitions from one potential well to the other, i.e., the fast component undergoes frequent reversals (red line in Fig. 2). In contrast, the slow particle, owing to its greater inertia, requires substantially more energy to cross between wells (blue line in Fig. 2). As a result, it often pulls the fast particle back to the original well, thereby restoring the initial polarity (as reflected in the similarity between the blue and black lines). A successful reversal of the axial component occurs only when both particles simultaneously approach the threshold of transition to the opposite well. This behaviour arises because the slow particle contributes most to the total axial component, as described by equation (5). Consequently, the model suggests that, prior to a reversal, both particles must drift toward the transition threshold—a process that takes time and reproduces the slow decay phase documented in palaeomagnetic records. Once the reversal occurs, both particles rapidly move toward the minimum of the opposite well. This swift transition reflects the fast recovery phase typically observed after a full dipolar field reversal. In this framework, reversal asymmetry emerges naturally from the elastic interaction between

the slow and fast particles. This feature is robust across numerous simulations and is not merely a statistical anomaly (see Fig. 3, panels E and F, in Molina-Cardín *et al*. 2021).

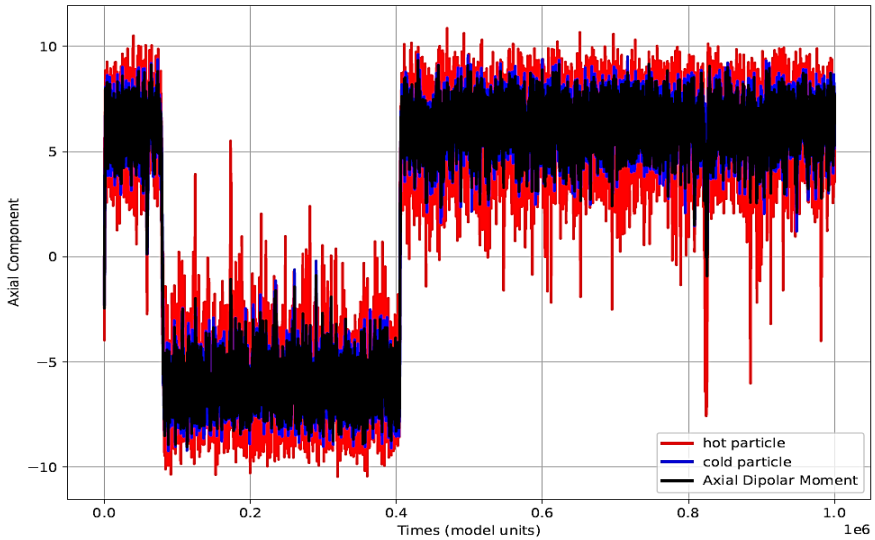


Fig. 2 Typical time series generated by the diffusive particles-in-a-well model (4). The x-axis is expressed in arbitrary time units determined by the characteristic scales of the variables involved in equations (3) and (4). The y-axis represents the axial component of the magnetic dipole, which is not constrained to absolute values of the Geomagnetic Dipole Moment. The term “hot particle” is used interchangeably with “fast particle,” and “cold particle” with “slow particle.” These terms are retained in the legend to allow readers to readily identify the analogous plots presented by Molina-Cardín *et al*. (2021). The axial component is calculated according to equation (5) with $w = 0.2$.

A very similar time series is generated by model (4) as well (plot not shown). However, a comparison of the probability density functions (PDFs) produced by models (3) and (4) reveals notable differences, particularly at the margins (see Fig. 3). These differences indicate that the models capture distinct behaviors during both low- and high-intensity periods—low-intensity phases typically occurring prior to and during reversals, and high-intensity phases characterizing stable epochs. Model (4), the diffusive model, tends to exhibit more frequent reversals and extended periods of stability. This suggests that the decay phase leading up to a reversal is longer, allowing the field to remain in a low-intensity

state for a greater duration. Conversely, the recovery phase following a reversal is much sharper, marked by a rapid transition to a high-intensity state. This interpretation is further supported by the dips observed around the values 5.0 and 15.0 in the PDF. Moreover, the central peak indicates that geomagnetic excursions are more frequent in model (4), likely due to the faster dynamics of the fast particle. In summary, compared to model (3), model (4) is characterized by slower decay, faster recovery, and a higher frequency of excursions.

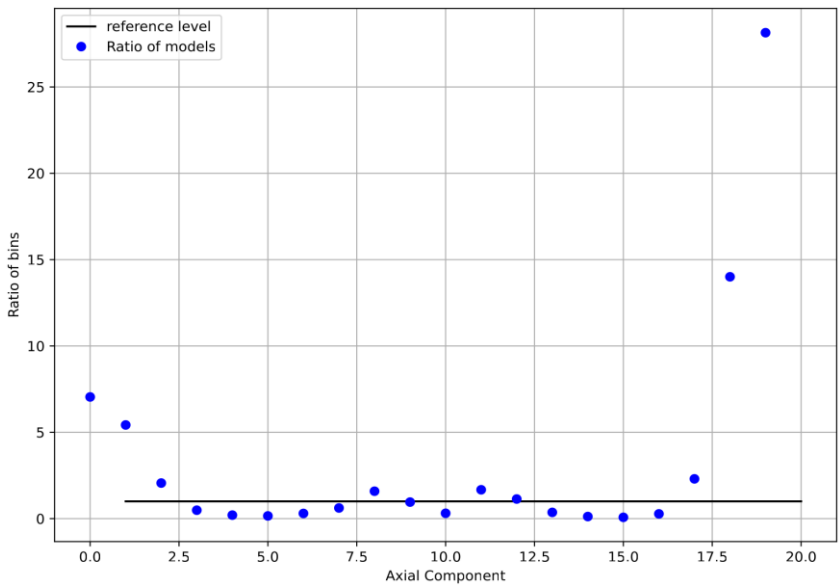


Fig. 3. Ratio plot of the PDFs of the Particles-in-well models (4) and (3). The horizontal line at 1 represents the case where the ratio of the corresponding bins is equal. In this plot, the PDF is divided into 20 bins, which should not be interpreted as absolute values of the axial component of the Dipole Moment. Increasing the resolution to 50 bins does not produce any noticeable qualitative changes, suggesting that the observed differences between the models are essential and robust.

A rather concerning issue arises when comparing the reversal patterns produced by the models with observational data. Figure 4 illustrates the reversal sequences generated by models (3) and (4), respectively. Both models produce polarity intervals of varying lengths; however, model (3) tends to yield longer intervals of stable polarity, indicating less frequent

reversals. This observation helps explain the differences seen in the right-hand bins of the ratio plot in Fig. 3, which reflect reversal statistics. Nevertheless, the total number of reversals generated in these simulations remains relatively low, even after extending the runs beyond 10^6 model time units. The limited reversal count prevents us from drawing firm conclusions as to whether either model (3) or model (4) qualitatively reproduces the reversal frequency observed in the palaeomagnetic record (cf. Fig. 1).

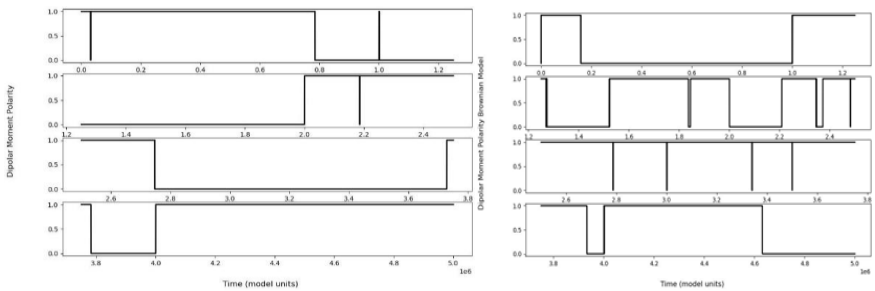


Fig. 4. Reversal patterns generated by the Langevin-type model (3) (left panel) and the Diffusive model (4) (right panel). The *y*-axis indicates the polarity of the dipolar field rather than its magnitude. For clarity, the left panel labels the *y*-axis as "Dipolar Moment Polarity," while the right panel labels it as "Dipolar Moment Polarity (Brownian Model)."

One might suspect that the small number of reversals is not due to insufficiently long simulations. The coefficients of the potential $V(x)$, as employed by Molina-Cardin *et al.* (2021), are $a = 0.00179$ and $b = -0.149$. Furthermore, these authors used a record of the last 800 kyr (Guyodo and Valet, 1999), which contains only one reversal—the Brunhes–Matuyama reversal. In contrast, here we use another record, namely the 4 Myr relative palaeointensity record of Valet and Meynadier (1993) which spans several reversals. The respective coefficients are: $a = 3.188$ and $b = -7.241$ To increase the probability of reversals, the magnitude of each random term in equations (3) or (4) must be significantly higher than that employed by Molina-Cardín *et al*. (2021). A typical simulation of the model with the new parameters yields an (unconstrained) axial component time series qualitatively similar to that shown in Fig. 2 (and is therefore not reproduced here). The change of coefficients does not alter the characteristic dynamics of the system. However, the situation differs

regarding the ratio plot of the two models (3) and (4). As shown in Fig. 5, model (4) dominates in the low and high magnitudes, whilst model (3) dominates in the middle region of the distribution (dominates in the sense that the ratio is large in the boundaries and quite small in the middle). This dependence indicates model (4) produces longer stable periods associated with strong dipolar field, whereas model (3) does not produce a typical dipole-dominated field. The domination of model (4) in the low end of the distribution (associated with very weak dipolar field that is typically observed during reversals) indicates that the particles in the diffusive model are so prone to reversals, as it is corroborated by the right panel in Fig. 6.

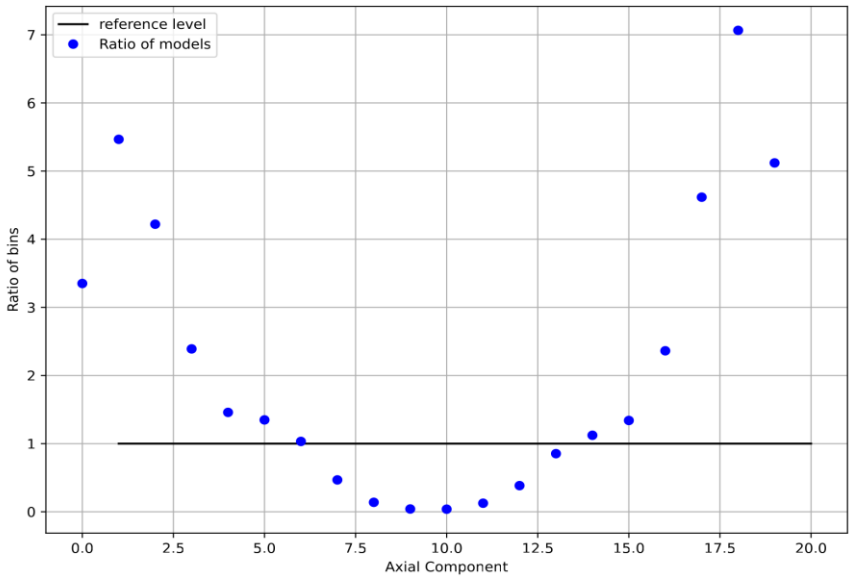


Fig. 5. Ratio plot of the PDFs of the particles-in-wells models (4) and (3), with coefficients determined from the 4 Myr record of Valet and Meynadier (1993). The horizontal line at value 1 corresponds to equal bin ratios. The PDF is constructed with 20 bins, which should not be interpreted as direct values of the axial component of the Dipole Moment. Increasing the number of bins to 50 produces no noticeable qualitative changes, reinforcing that the observed differences between the models are essential rather than numerical artifacts.

The ratio plot reveals that model (4) dominates at the edges of the axial component distribution, whereas model (3) dominates in the central bins. This indicates that model (4) spends more time in states

corresponding to strong dipolar configurations—i.e., near stable polarity states—while model (3) exhibits a higher probability density near the center, corresponding to low dipole amplitudes typically associated with transitional states such as reversals or excursions. Consequently, model (3) appears more prone to generating frequent or prolonged excursions and reversals, reflecting a more dynamic or less stable magnetic regime. By contrast, model (4) favors long-term stability in the axial dipole, with fewer and potentially sharper transitions between polarity states. These distinct occupancy patterns across the axial component range highlight fundamental differences in reversal dynamics between the two models, which may be traced back to differences in damping, inertia, and stochastic forcing structures.

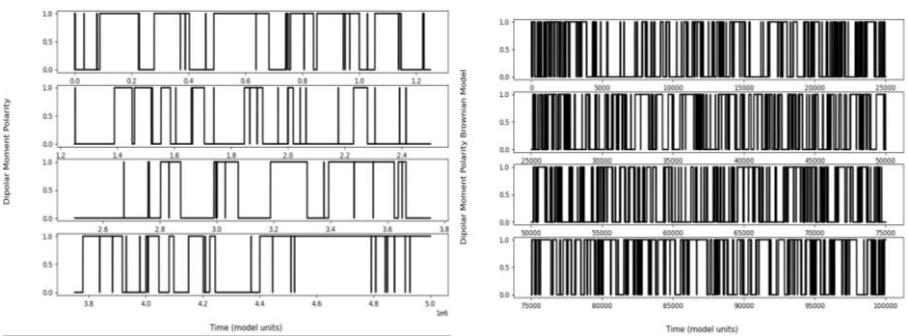


Fig. 6. Reversal patterns generated by the Langevin-type model (3) (left) and the diffusive model (4) (right), using coefficients estimated from the 4 Myr record of Valet and Meynadier (1993). The *y*-axis indicates polarity, not the magnitude of the dipolar field. In the left panel, the axis is labeled "Dipolar Moment Polarity"; in the right panel, "Dipolar Moment Polarity Brownian Model."

The polarity plots obtained by fitting the model coefficients to geomagnetic time series that include multiple reversals highlight the sensitivity of the particle-in-a-well framework to the particular record employed. For example, when applied to the classical Cande and Kent (1992) polarity timescale, the diffusive model generates an unrealistically high frequency of rapid reversals—an artifact inconsistent with paleomagnetic observations (Cande and Kent, 1992; 1995). In contrast, the Langevin (second-order) variant produces reversal sequences that align more closely with the rock record: the spacing, clustering, and occurrence of extended "quiet" intervals between reversals resemble empirical patterns with greater fidelity. Although the present study illustrates results

primarily from the Cande and Kent (1992) dataset, we have repeated the analysis using extended and updated datasets, including the Cande and Kent (1995) timescale, updated Geomagnetic Polarity Timescale (GPTS) compilations for the Cenozoic and Neogene, and sedimentary/igneous records calibrated with10/Be and paleointensity reconstructions (Singer *et al*. 2019). In all cases, the qualitative outcome remains consistent: the diffusive model systematically overproduces rapid reversals, while the Langevin formulation yields reversal sequences with clustering and temporal structure that are more realistic. Supplemental figures are omitted here, but a full quantitative comparison will be presented in a forthcoming publication.

4. DISCUSSIONS

The particles-in-a-well model, though simplified and heuristic, remains a valuable framework for capturing key statistical features of geomagnetic dipole reversals—particularly the asymmetry between decay and recovery phases. Nevertheless, recent analyses demonstrate that the model's behavior is highly sensitive to the choice of palaeomagnetic record used to construct the underlying potential $V(x)$. Earlier formulations, based on the Brunhes–Matuyama reversal recorded over the past 800 kyr (Guyodo and Valet, 1999; 2006), provided limited temporal coverage, encompassing only a single reversal event. To overcome this limitation, we employ more extensive datasets, such as the 4 Myr record of Valet and Meynadier (1993), which captures multiple reversal events. This extended record reflects essential long-term geomagnetic behavior, including intervals of both high and low intensity. However, while it improves the statistical robustness of the model, it does not fully account for very long timescales spanning tens of millions of years, which remain beyond the scope of the present analysis.

Crucially, our findings demonstrate that models derived from extended records behave differently: while the diffusive model tends to overproduce reversals in rapid succession, the Langevin model generates polarity evolutions that more closely match the empirical record. This reinforces earlier arguments (Valet and Meynadier, 1993; Valet *et al*. 2005) regarding the necessity of comprehensive reversal datasets for accurately modelling geomagnetic behaviour. A broader temporal dataset not only enhances the model's capacity to reproduce diverse geodynamo regimes but also improves its statistical robustness, reducing uncertainty

and allowing for more reliable differentiation between genuine geomagnetic signals and random noise (Backus *et al.* 1996; Olson *et al.* 2010). Moreover, the inclusion of extended records increases the likelihood of capturing rare but significant geomagnetic phenomena, such as excursions. Nevertheless, as highlighted by McMillan and Constable (2006) and Panovska *et al*. (2015), the use of older palaeomagnetic records must be approached with caution due to potential issues of temporal resolution, data quality, and post-depositional alteration. Rigorous calibration and cross-validation remain essential before such datasets can reliably inform model construction and interpretation.

The fast and slow particles in the model serve as conceptual proxies for distinct dynamical components of the geomagnetic field, each evolving on characteristic timescales. While these particles do not correspond directly to specific physical structures in the Earth's outer core, their behaviour reflects mechanisms identified in high-fidelity geodynamo simulations. It is well established that the geodynamo operates across a broad spectrum of temporal and spatial scales, driven by turbulent convection in the electrically conducting outer core and modulated by interactions with the solid inner core and mantle (Hollerbach and Jones, 1993; Hollerbach and Jones, 1995; Backus *et al*. 1996; Davidson, 2013). In this context, the slow particle may be interpreted as representing the long-term evolution of the axial dipole field, which numerical simulations demonstrate is stabilized by large-scale, columnar convection aligned with the rotation axis (Aubert *et al*. 2017). These simulations—particularly in the asymptotic rapidly rotating regime—highlight quasi-geostrophic flows that maintain a persistent dipole-dominated magnetic field, analogous to the slowly evolving component captured by the inertial particle.

Conversely, the fast particle may be understood as representing more transient, localized, or non-dipolar contributions to the field, driven by rapidly fluctuating small-scale convection or episodic interactions between convective upwellings and magnetic structures. These processes, observed in simulations, can induce short-term variability and, under certain conditions, trigger reversals (Dormy and Soward, 2007). In addition, mantle control on core dynamics—through thermal anomalies or compositional heterogeneities at the core–mantle boundary—can impose long-wavelength, low-frequency forcing that modulates reversal frequency and stability, including the occurrence of superchrons (Olson and Amit, 2014). While the low-dimensional model does not resolve such spatial structures explicitly, its effective separation into slow and fast

modes reflects this multiscale architecture. Thus, the two-particle system can be viewed as a minimal representation of the competing influences within the dynamo: one stabilizing and long-lived, the other destabilizing and reactive, together shaping the reversal statistics observed in both simulations and the palaeomagnetic record.

The mechanism of geomagnetic reversals in the particles-in-a-well model hinges on the elastic interaction between the two particles, represented by the coupling spring. We suspect that this elastic interaction is a key element of the reversal process. Specifically, the coupling may serve as a mechanism for the gradual accumulation of energy, producing the slow decay phase, followed by a rapid release during the recovery phase after a reversal. This dynamic mirrors the asymmetry observed in geomagnetic reversals. We propose that elastic interactions offer a useful conceptual framework for studying energy transfer processes within the geodynamo. By modeling how energy is stored and subsequently released, these interactions could help illuminate the physical mechanisms underlying geomagnetic field reversals.

5. CONCLUSIONS

The particles-in-a-well model provides a minimal yet insightful framework for exploring the statistical and dynamical features of geomagnetic field reversals. Within this model, reversals are interpreted as stochastic transitions between two stable states of the magnetic dipole, triggered by fluctuations strong enough to overcome an energy barrier. This approach successfully reproduces key observational features of the reversal process, particularly its intermittent nature—characterized by long periods of stable polarity punctuated by rapid transitions—as well as the asymmetry often observed between the decay and recovery phases of the dipole field. A central component of the model's success lies in the inclusion of an elastic coupling between two particles evolving on different timescales, which facilitates energy exchange and controls the reversal dynamics in a physically interpretable manner.

In this study, we have shown that the diffusive and Langevin versions of the model yield markedly different statistical behaviors. When calibrated against extended geomagnetic polarity timescales (Cande and Kent, 1992; Cande and Kent, 1995; Singer *et al*. 2019), the Langevin model proves more consistent with palaeomagnetic observations, avoiding the unrealistic clustering of reversals seen in the diffusive version. This

finding highlights the importance of incorporating second-order dynamics to realistically suppress spurious reversal events and better reflect the stability observed in the long-term geomagnetic record. Moreover, the model's sensitivity to the choice of input dataset underscores the need for using high-quality, well-resolved reversal records when constructing effective potential functions.

Despite these promising features, the two-particle system remains a highly simplified abstraction of the geodynamo. Reducing the complex, multiscale dynamics of Earth's outer core to only two degrees of freedom inevitably limits the model's ability to capture the full range of geomagnetic behavior, particularly when it comes to rare events such as excursions, superchrons, or the influence of mantle heterogeneities. Building on the idea of elastic interactions between dynamical modes, future work will explore their incorporation into more elaborate frameworks, such as the domino model (Duka *et al*. 2015; Peqini *et al*. 2015; Peqini, 2024), which offers greater potential for capturing the spatial and temporal richness of the Earth's magnetic field. This next step aims to bridge the gap between low-dimensional models and fully 3D geodynamo simulations, providing a more comprehensive yet computationally tractable description of geomagnetic reversals.

ACKNOWLEDGEMENTS

We gratefully acknowledge the use of the HPC Cluster acquired through the project *Ngritja e një qendre llogaritëse në mbështetje të bashkëpunimit të IAL-ve shqiptare me projektin europian Compact Muon Solenoid (CMS) në CERN*, funded by AKKSHI. The authors declare no competing interests related to the content of this paper. Further information regarding the algorithms and codes can be found in Peqini *et al*. (2015), where the codes developed for another model can be readily adapted for the particles-in-a-well framework. For additional technical details, interested readers are encouraged to contact the corresponding author directly.

Data accessibility:

There are no databases to be made public, but the codes are available upon request to the author.

Declaration of AI use: There has been no use of AI when writing the actual paper.

Authors' contributions:

K.P.: conceptualization, writing—original draft; validation, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration: The author declares there are no conflicts of interest.

Funding: The author declares no funding was received when writing this paper.

REFERENCES

Aubert J, Finlay CC, Fournier A. 2013. Bottom-up control of geomagnetic secular variation by the Earth's inner core. *Nature,* **502 (7470):** 219–223. PMID: 24108054. DOI: 10.1038/nature12574.

Aubert J, Gastine T, Fournier A. 2017. Spherical convective dynamos in the rapidly rotating asymptotic regime. *Journal of Fluid Mechanics*, **813**, 558–593. DOI:10.1017/jfm.2016.789.

Backus G, Constable C, Parker R. 1996. *Foundations of Geomagnetism*. New York, NY: Cambridge University Press.

Buffett BA, Ziegler L, Constable CG. 2013. A stochastic model for palaeomagnetic field variations. *Geophysics Journal International***, 195: 86–97.** doi: 10.1093/gji/ggt218.

Cande SC, Kent DV. 1992. A new geomagnetic polarity time scale for the late Cretaceous and Cenozoic. *Journal of Geophysical Research*, **97 (B10):** 13917–13951. https://doi.org/10.1029/92JB01202.

Cande SC, Kent DV. 1995. Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *Journal of Geophysical Research*, **100 (B4):** 6093–6095. https://doi.org/10.1029/94JB03098.

Constable C, Morzfeld M. 2025. Weather at the core: defining and categorizing geomagnetic excursions and reversals. *Geophysical Journal International*, **240 (1):** 747–762**.** https://doi.org/10.1093/gji/ggae415.

Davidson PA, 2013a. Scaling laws for planetary dynamos. *Geophysical Journal International*, **195 (1):** 67-74. https://doi.org/10.1093/gji/ggt227.

Davidson PA. 2013b. Turbulence in rotating, stratified and electrically conducting fluids. Cambridge University Press. Online ISBN 781139208673 DOI: https://doi.org/10.1017/CBO9781139208673.

Dormy E, Soward AM. 2007. *Mathematical Aspects of Natural Dynamos*. New York: CRC Press, Taylor and Francis Group. Online ISBN: 9780429145643. https://doi.org/10.1201/9781420055269.

Duka B, Peqini K, De Santis A, Francisco-Carrasco JP. 2015. Using domino model to study the secular variation of the geomagnetic dipolar moment. *Physics of Earth and Planetary Interiors*, **242**: 9-23.http://dx.doi.org/10.1016/j.pepi.2015.03.001.

Gubbins D. 1988. *Mechanism for geomagnetic polarity reversals*. Nature, **335**: 525–528. 10.1038/326167a0

Gubbins D. 1999. The distinction between geomagnetic excursions and reversals. *Geophysical Journal International*, **137(1):** F1–F4. https://doi.org/10.1046/j.1365-246x.1999.00810.x

Guyodo Y, Valet JP. 1999. Global changes in intensity of the Earth's magnetic field during the past 800 kyr. *Nature,* **399**: 249–252. https://doi.org/10.1038/20420.

Guyodo Y, Valet JP. 2006. A comparison of relative paleointensity records of the Matuyama Chron for the period 0.75-1.25 Ma. *Physics of the Earth and Planetary Interiors*, **156 (3-4):** 205–212. https://doi.org/10.1016/j.pepi.2005.03.020.

Hollerbach R, Jones CA. 1993. Influence of the Earth's inner core on geomagnetic fluctuations and reversals. *Nature*, **365**: 541–543. https://doi.org/10.1038/365541a0.

Hollerbach R, Jones CA. 1995. On the magnetically stabilizing role of the Earth's inner core. *Physics of the Earth and Planetary Interiors*, **87 (3-4):** 171–181. https://doi.org/10.1016/0031-9201(94)02965-E.

Hoyng P, Schmitt D, Ossendrijver M. 2002. A theoretical analysis of the observed variability of the geomagnetic dipole field. *Physics of the Earth and Planetary Interiors*, **130**: 143–157. https://doi.org/10.1016/S0031-9201(02)00004-3.

Kageyama A, Sato T. 1997. Velocity and magnetic field structures in a magnetohydrodynamic dynamo. *Physics of Plasmas*, **4 (5):** 1569–1575. https://doi.org/10.1063/1.872287.

McMillan DG, Constable CG. 2006. Limitations in correlation of regional relative geomagnetic paleointensity. *Geochemistry, Geophysics, Geosystems*, **7 (9):** Q09009. https://doi.org/10.1029/2006GC001350.

Merrill RT, McElhinny MW, McFadden PL. 1996. The magnetic field of the Earth: Paleomagnetism, the core, and the deep mantle. International geophysics series. Volume **63**. Academic Press, San Diego. ISBN: 0-12-491246-X.

Molina-Cardín A, Dinis L, Osete ML. 2021. Simple stochastic model for geomagnetic excursions and reversals reproduces the temporal asymmetry of the axial dipole moment. *Earth, Atmospheric, and Planetary Sciences*, **118(10)**: E2017696118**.** https://doi.org/10.1073/pnas.2017696118.

Mori N, Schmitt D, Wicht J, Ferriz-Mas A, Mouri H, Nakamichi A, and Morikawa M. 2013. Domino model for geomagnetic field reversals. *Physical Review E* **87**: 012108. DOI: https://doi.org/10.1103/PhysRevE.87.012108.

Morzfeld M, Buffett BA. 2019. A comprehensive model for the kyr and Myr timescales of Earth's axial magnetic dipole field. *Nonlinear Processes in Geophysics*, **26 (3)**: 123–142. https://doi.org/10.5194/npg-26-123-2019.

Olson P, Amit H. 2014. *Mantle superplumes induce geomagnetic superchrons*. Frontiers in Earth Science, **2**, 22. https://doi.org/10.3389/feart.2015.00038.

Olson PL, Coe RS, Driscoll PE, Glatzmaier GA, Roberts PH, 2010. Geodynamo reversal frequency and heterogeneous core-mantle boundary heat flow. *Physics of Earth and Planetary Interiors*, **180 (1-2):** 66–79. 10.1016/j.pepi.2010.02.010.

Panovska S, Korte M, Finlay CC, Constable CG. 2015. Limitations in paleomagnetic data and modelling techniques and their impact on Holocene geomagnetic field models. *Geophysical Journal International*, **202 (1):** 402–418. https://doi.org/10.1093/gji/ggv137.

Peqini K, Duka B, De Santis A. 2015. Insights into pre-reversal paleosecular variation from stochastic models. *Frontiers in Earth Science*, **3 (52):** 1-13. https://doi.org/10.3389/feart.2015.00052.

Peqini K, Koçi E, Prenga D, Osmanaj R. 2023. The core-mantle boundary velocity field in the recent decades. *AIP Conference Proceedings*, **2872**: 120063. https://doi.org/10.1063/5.0162927.

Peqini K. 2024. Inertial forces reproduce the observed variations of the frequency of geomagnetic reversals. *Albanian Journal of Natural and Technical Sciences* (AJNTS), **61 (2):**3-18. https://akad.gov.al/wp-content/uploads/2025/06/1-Peqini-K.-Inertial-forces-reproduce-the-observed-variations-_Albanian-Journal-of-Natural-and-Technical-Sciences-20242-XXIX-61.pdf.

Schmitt DM, Ossendrijver MAJH, Hoyng P. 2001. Magnetic field reversals and secular variation in a bistable geodynamo model. *Physics of the Earth and Planetary Interiors*, **125 (1-4):** 119–124. https://doi.org/10.1016/S0031-9201(01)00237-0.

Singer BS, Jicha BR, Mochizuki N, Coe RS. 2019. Synchronizing volcanic, sedimentary, and ice core records of Earth's last magnetic polarity reversal. *Science Advances*, **5 (8):** eaaw4621. doi: 10.1126/sciadv.aaw4621.

Valet JP, Fournier A. 2016. Deciphering records of geomagnetic reversals. *Reviews of geophysics*, **54 (2)**: 410–446. https://doi.org/10.1002/2015RG000506.

Valet JP, Meynadier L. 1993. Geomagnetic field intensity and reversals during the past four million years. *Nature*, **366**: 234–238. 10.1038/366234a0.

Valet JP, Meynadier M, Guyodo Y. 2005. Geomagnetic dipole strength and reversal rate over the past two million years. *Nature*, **435**:802–805. doi: 10.1038/nature03674.

Wicht J, Meduri D. 2016. A Gaussian model for simulated geomagnetic field reversals. Physics of the Earth and Planetary Interiors. **259**: 45–60. https://doi.org/10.1016/j.pepi.2016.07.007.

Ziegler LB, Constable CG. 2011. Asymmetry in growth and decay of the geomagnetic dipole. Earth and Planetary Science Letters, **312 (3–4):** 300-304. https://doi.org/10.1016/j.epsl.2011.10.019.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

NEAR REAL-TIME MOMENT TENSOR PROCESSING AND VALIDATION IN THE ADRIA–ALBANIAN OROGENIC CONTACT: AN OPERATIONAL FRAMEWORK ON THE ALBANIAN BROADBAND SEISMIC NETWORK

Besian RAMA, Edmond DUSHI and Damiano KOXHAJ

Institute of Geosciences (IGEO), Polytechnic University of Tirana, Albania
Author for correspondence: b.rama@geo.edu.al
BR, 0000-0002-6759-2130 ; ED, 0000-0001-7599-7034;
DK, 0009-0002-9792-9010
https://doi.org/10.67537/ajnts3001002

ABSTRACT

The expansion of seismic station coverage and the continuous acquisition of broadband waveform data in Albania—particularly following the significant earthquake of November 2019—have underscored the need for an automated, scalable, and reliable workflow for rapid seismic source characterization. In response, an operational framework has been developed to enable near real-time processing and validation of seismic data and moment tensor (MT) solutions across the tectonic boundary between the Adria microplate and the Albanides, where complex deformation processes define the orogenic system. The framework operates on the Albanian Broadband Seismic Network and is complemented by regional broadband stations along the eastern Adriatic margin. It integrates automated modules for event detection and location using SeisComP 5.4, waveform conditioning, near real-time MT inversion with GISOLA, and subsequent validation and focal mechanism clustering through the FMC algorithm. Since 2014, the system has generated over 1,300 MT solutions, with moment magnitudes (Mw) ranging from 2.5 to 6.4. Iterative revision cycles have progressively enhanced correlation and variance-reduction metrics. Comprehensive quality assessment of MT results, combined with continuous dissemination of validated solutions to authoritative seismological institutions such as the Euro-Mediterranean Seismological Centre (EMSC), attests to the robustness, reliability, and interoperability of the implemented procedures. This presentation highlights the modular architecture, validation protocols, and analytical performance of the operational MT framework, which now constitutes

a core component of real-time seismic monitoring and seismotectonic interpretation in Albania.

***Keywords*:** Near real-time moment tensor; SeisComP; GISOLA; Albanian Broadband Seismic Network; Adria–Albanian orogenic contact; operational seismology

1. INTRODUCTION

Albania is situated at the active collision front between the Adriatic (Adria) and Eurasian plates, where crustal shortening across the external thrust belt coexists with strike-slip and localized extensional deformation in the hinterland (Aliaj, 2006). This complex tectonic configuration results in frequent moderate-magnitude earthquakes and occasional damaging events (Aliaj *et al*. 2010; Aliaj, 2020), emphasizing the necessity for robust focal mechanism and moment tensor (MT) information to effectively characterize seismogenic zones and faulting processes. Such data are critical for the prompt association of earthquakes with their causative structures and for the development of physically consistent models supporting future seismic-hazard assessments.

The Mw 6.4 Durrës earthquake of 26 November 2019—a blind thrust-faulting event near the Adria–Eurasia boundary—was the strongest earthquake in Albania in several decades, causing extensive damage and casualties across the western part of the country (Moshou *et al*. 2019). This event underscored the role of deep-seated thrusts in regional seismogenesis, complementing the shallow compressional faulting typically associated with the Outer Albanides (Aliaj *et al*. 2010; Teloni *et al*. 2021; Matraku *et al*. 2023;). Depth estimates ranging between 20 and 40 km, reported by EMSC**,** USGS**,** GFZ**,** and IGEO (Fig. 1), challenged earlier interpretations that characterized Albanian seismicity as predominantly shallow. At the time, the limited number of operational seismic stations (Fig. 1), discrepancies in focal depth and mechanism determinations among international agencies, and a partially functional accelerometric network restricted near-field recordings and contributed to large source-parameter uncertainties. The scarcity of broadband sensors west of the internal–external structural boundary (Fig. 1) further highlighted the need for a denser, automated, and interoperable monitoring system capable of managing the rapidly increasing volume of waveform data in near real time.

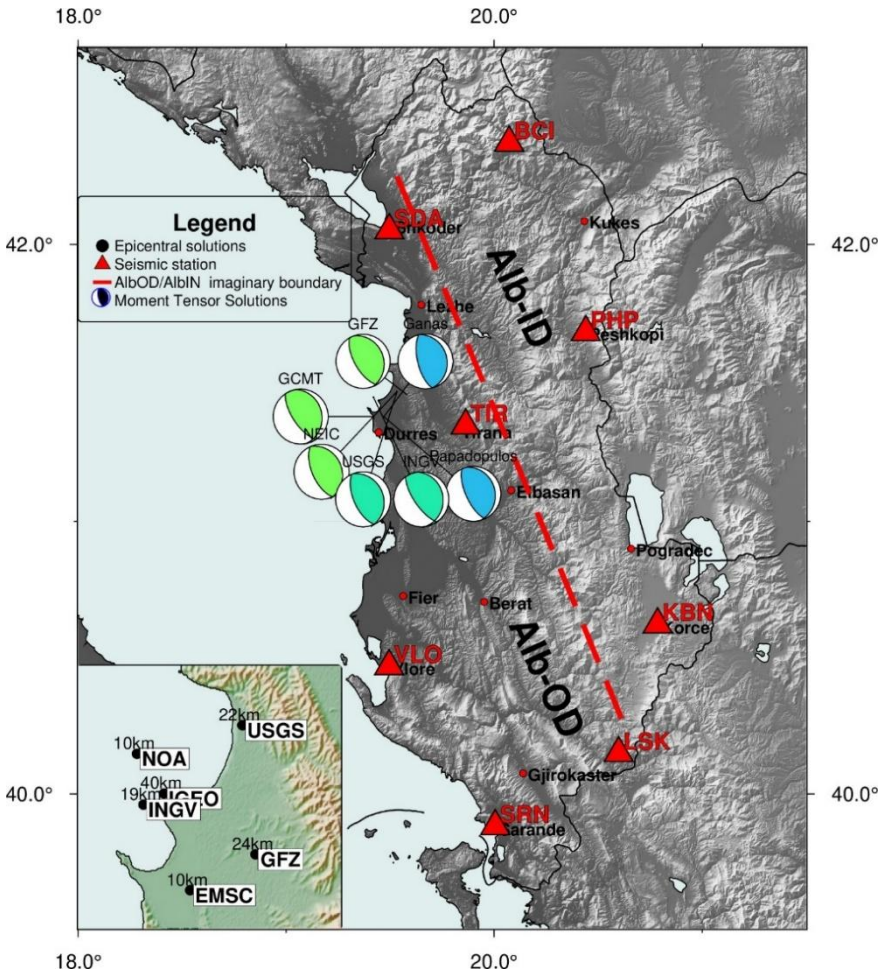


Fig. 1. Epicentral locations and moment-tensor (MT) solutions of the 26 November 2019 Durrës earthquake sequence from major seismological agencies (EMSC, USGS, GFZ, INGV, NOA, GCMT, IGEO). The red dashed line marks the indicative boundary between the Outer and Inner Albanides, while the broadband stations active at the time show the sparse network coverage before densification (Dushi *et al*. 2025). Figure generated using Generic Mapping Tools (GMT 6.0.0; Wessel, 2019).

Since 2020, several initiatives have significantly enhanced Albania's seismic observation capacity. The RESEAL Project (UNDP, 2024) expanded the permanent broadband network, while the multinational

AdriaArray framework established a coordinated infrastructure for dense, interoperable broadband deployments across the Adriatic region. Within this collaborative context, the Institute of Geosciences (IGEO) has played a leading role in station deployment, metadata coordination, and data integration (Dushi *et al*. 2025, published online October 2025). These developments provided the foundation for the establishment of an operational system for near real-time moment tensor (MT) processing and validation, designed to capture the tectonic complexity of the Adria–Albanian orogenic contact.

The presented near real-time operational chain, conceived as a software-engineered infrastructure, is described in detail in the following sections, with emphasis on its two principal modules: *SeisComP 5.4*, responsible for event detection, cataloging, and data logistics; and GISOLA (Triantafyllis, 2021), dedicated to MT inversion and validation. The applied quality-control procedures are discussed both from operational and scientific perspectives, highlighting their role in ensuring reproducible MT solutions with quantified quality metrics, refining results through iterative revision cycles, and synthesizing spatial faulting patterns for seismotectonic interpretation and hazard assessment.

2. Data Source and Quality

The waveform data used for the near real-time MT inversions are acquired through the Albanian Seismological Network (ASN), operated by the Department of Seismology at the Institute of Geosciences (IGEO). All real-time acquisition, quality control, and data management activities are conducted by the National Center for Seismic Monitoring (NCSM)**,** which also facilitates the dissemination of both raw and processed data via national and international data-exchange channels (Institute of Geosciences, 2023).

At the time of the Mw 6.4 Durrës earthquake on 26 November 2019, the national network comprised 8 broadband (BB) and 17 accelerometric (SM) stations (Fig. 2a). Although this configuration was sufficient for basic earthquake location and intensity mapping, it exhibited substantial azimuthal gaps—particularly along the eastern Adriatic margin—which limited both the stability and resolution of regional MT inversions for events below approximately M 3.0–3.5.

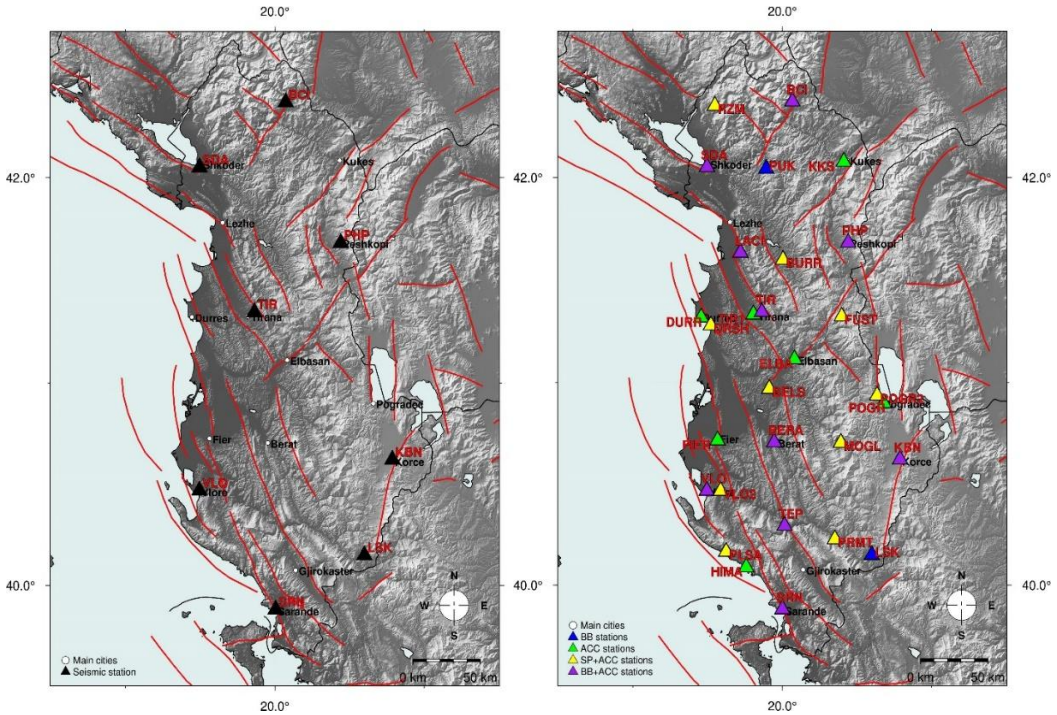


Fig. 2. (a) Albanian Seismic Network operating during the 2019 Durrës earthquake sequence; (b) modernized configuration in 2025 showing broadband (BB), accelerometric (ACC), short-period (SP), and combined BB–ACC stations. Figure generated using Generic Mapping Tools (GMT 6.0.0; Wessel, 2019).

Following the 2019 Durrës earthquake, a series of national and regional initiatives markedly enhanced Albania's seismic network coverage, station reliability, and waveform quality. Between 2020 and 2022, the first strengthening phase prioritized service continuity and data logistics over simple spatial expansion. Two new broadband stations were installed, telemetry was standardized under SeedLink, and inverter–battery backup systems were deployed at key nodes, significantly reducing data latency and packet loss. Continuous data streaming into a standardized SeisComP Data Structure (SDS) archive ensured compatibility with both near real-time and offline processing pipelines. In parallel, the Albanian Seismological Network (ASN) achieved full integration within the

European FDSN framework (network code AC; Institute of Geosciences, 2002) and became routinely accessible through the European Integrated Data Archive (EIDA) hosted by INGV (Italy), enabling real-time cross-border data exchange

In late 2022, five Streckeisen STS-2 broadband sensors were installed at Leskovik (LSK), Peshkopi (PHP), Berat (BERA), Laç (LACI), and Tepelenë (TEP). These installations substantially improved low-frequency fidelity and signal-to-noise ratio (SNR) along regional paths—parameters essential for stable MT inversions and consistent variance-reduction performance. During 2023–2024, network densification accelerated through the RESEAL Project (UNDP / Government of Albania), which established ten new permanent sites equipped with SS10 1 Hz short-period sensors and collocated accelerometers recorded on SARA SL06C6 digitizers. Existing nodes were upgraded with higher-resolution digitizers, enhanced telemetry, and solar-powered backup systems. The accelerometric network expanded to 27 stations, strategically distributed across both bedrock and soft-soil environments in major urban centers to ensure representative recording of site-specific ground motions. The resulting configuration (Fig. 2b) represents the first dense broadband–accelerometric hybrid system managed under IGEO's operational control.

By 2025, the ASN comprised 13 broadband, 9 short-period, and 27 strong-motion stations, supplemented by portable units for temporary deployments. Data streams are continuously acquired using SeisComP 5.4 via SeedLink, archived in SDS format, and made available through local FDSNWS services and EIDA. Beyond national coverage, the operational virtual network incorporates regional data from AdriaArray partners and cooperating networks, including the Mediterranean Very Broadband Seismographic Network (MN), the Italian National Seismic Network (IV) (Suárez, 2008), the Hellenic Unified Seismic Network (HL + HT) (Evangelidis, 2021), and networks from Montenegro (ME), Serbia (SJ), and Türkiye (TU) (Suárez, 2008). This configuration provides continuous broadband coverage across Albania and its tectonic boundaries, markedly improving azimuthal distribution and overall data redundancy (Fig. 3).

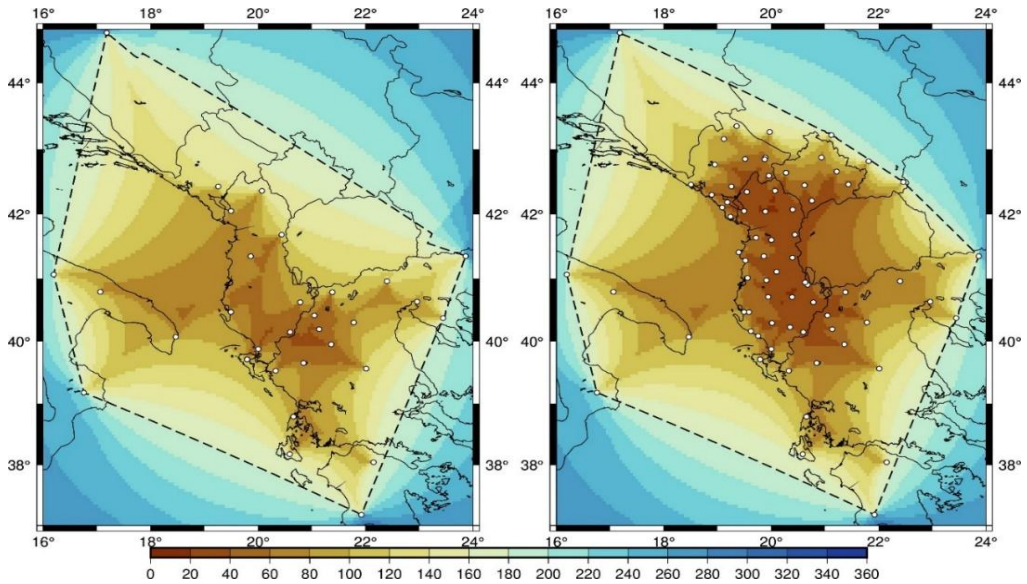


Fig. 3. Improvement of spatial and azimuthal coverage between 2019 (left) and 2025 (right), including partner-network stations forming the Albanian Seismic Virtual Network. Figure generated using Generic Mapping Tools (GMT 6.0.0; Wessel, 2019).

The cumulative effect of these upgrades has resulted in a substantial improvement in data quality**,** network geometry, and catalog completeness. Median azimuthal gaps decreased by approximately 10–25° across most magnitude ranges (Table 1), with the largest gains observed for events below M 3.0, where MT stability is particularly sensitive to geometric constraints. This systematic reduction in azimuthal gaps (Fig. 4) enhances the uniformity of station coverage, improves phase-arrival balance in location algorithms, and provides a more reliable dataset for moment tensor inversions**.**

Table 1. Azimuthal-gap reduction for events detected by the Albanian Seismological Network at NCSM between 2019 and 2024.

Magnitude bin (ML)	2019 gap (°)	2024 gap (°)	Gap reduction (°)
0.5–1.0	234.5	186	−48.5
1.0–1.5	180.5	174	−6.5
1.5–2.0	165	145	−20
2.0–2.5	140	122	−18

2.5–3.0	120	107	−13
3.0–3.5	115	105	−10
3.5–4.0	115.5	112	−3.5
4.0–4.5	127	103	−24
4.5–5.0	112	214	+102

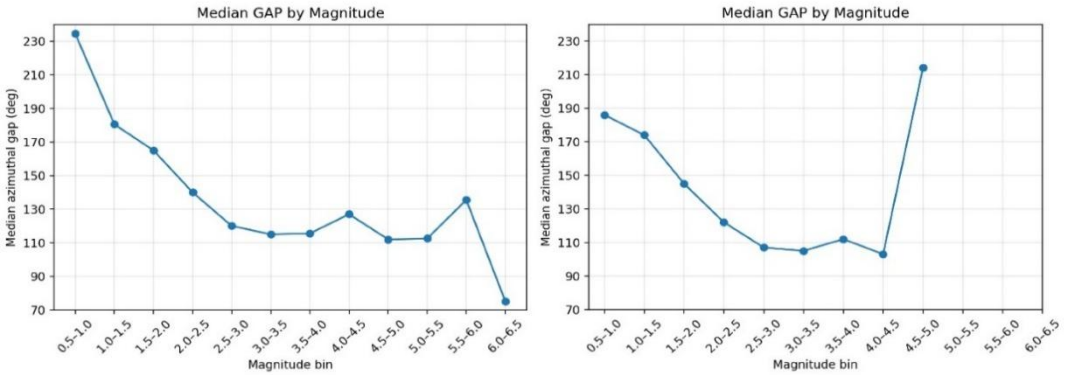


Fig.4: Median azimuthal gap versus magnitude bin showing systematic reduction for all magnitudes from 2019 (blue) to 2024 (orange), except for a sparse outlier due to limited sample size.

Network densification also produced a marked improvement in the magnitude of completeness (Mc), which decreased from approximately Mc = 2.25 in 2019 to Mc = 1.35 in 2024. This reduction reflects increased station density, broader bandwidth coverage, and improved telemetry uptime. The resulting shift enhances catalog completeness (Fig. 5), stabilizes b-value estimation in the Gutenberg–Richter distribution, and extends moment tensor (MT) applicability to smaller-magnitude events.

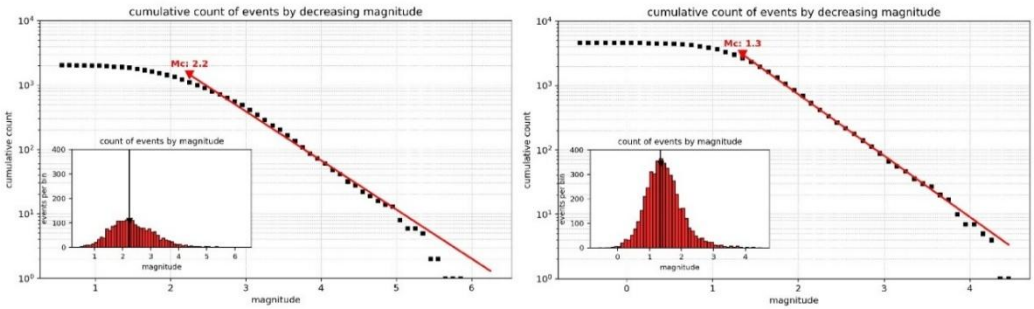


Fig. 5. Magnitude-of-completeness (Mc) improvement between 2019 and 2024 for the IGEO seismic catalog, demonstrating increased detection capacity for lower-magnitude events.

These quantitative gains in network geometry and completeness directly strengthen the near real-time MT inversion workflow, enabling more stable variance-reduction metrics and higher-confidence solutions across the Adria–Albanian contact.

The unified broadband**,** short-period, and accelerometric infrastructure developed through these initiatives constitutes a key prerequisite for near real-time MT inversion using GISOLA. Enhanced azimuthal coverage, reduced Mc, and improved waveform fidelity have transitioned Albania's seismic monitoring paradigm from a location-centric to a source-mechanism-centric operational model—an essential foundation for reliable real-time MT inversion and stress-field analysis.

3. System Architecture and Operational Workflow

The near real-time Moment Tensor (MT) processing system has been designed as an integrated, modular framework that couples SeisComP 5.4 with GISOLA within the operational environment of the National Center for Seismic Monitoring (NCSM) at the Institute of Geosciences (IGEO)**.** The architecture was developed to ensure high reliability**,** deterministic latency**,** and reproducibility of MT solutions, thereby meeting both operational and research requirements for seismic monitoring in Albania. SeisComP 5.4 (Gempa GmbH, 2008) manages waveform acquisition, phase picking, and hypocentral determination, while GISOLA performs automated MT inversion and validation, triggered directly by the SeisComP event stream (Fig. 6). The overall system operates as a software-engineered infrastructure designed to minimize manual intervention while producing auditable, reproducible, and scientifically traceable results. All processing modules write to a dedicated log-file system that records configuration parameters, station metadata, velocity models, inversion settings, and performance statistics—ensuring full traceability and reproducibility of each published solution.

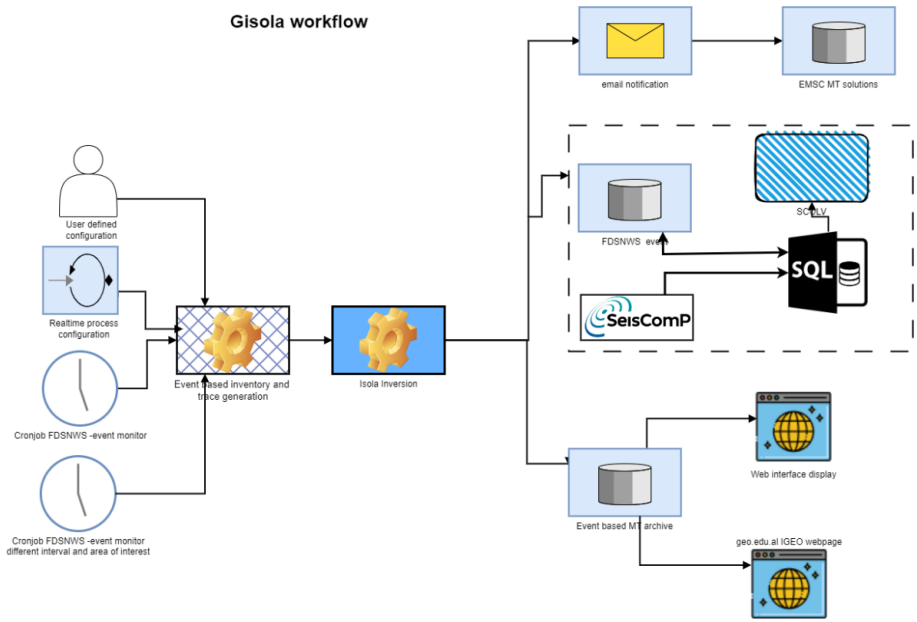


Fig. 6. Simplified architecture of the SeisComP–GISOLA operational chain implemented at the National Center for Seismic Monitoring, showing event triggering, data routing, inversion, and publication workflow.

The system triggering sequence begins with an automatic event detection by SeisComP, which locates the earthquake and publishes preliminary origin parameters via its internal web service, fully compliant with Federation of Digital Seismograph Networks (FDSN) standards. A watcher process continuously monitors this service at fixed intervals (every 540 **s**) and initiates a moment tensor (MT) inversion when the event origin meets predefined stability criteria in time, depth, magnitude, and hypocentral coordinates, or when the maximum waiting time is reached. This approach ensures deterministic latency while preventing redundant recalculations, thereby preserving the operational continuity required for near real-time processing**.**

SeisComP simultaneously provides waveform routing and metadata through its FDSNWS-station and FDSNWS-dataselect services, supplying GISOLA with synchronized waveform and event information.

Station metadata are dynamically prioritized according to event magnitude: small and moderate events are processed using higher-quality, low-noise stations, whereas larger events automatically expand the spatial

aperture to enhance azimuthal coverage without compromising inversion stability. Automated gating excludes marginal or inconsistent stations, typically yielding a geometry of six to sixteen contributing sites, depending on the signal-to-noise ratio and network distribution across the Adria–Albanian contact zone.

Waveform retrieval follows a hierarchical approach: the local SeisComP Data Structure (SDS) archive is queried first for rapid and reliable access, while FDSNWS-dataselect or SeedLink protocols serve as fallbacks when local data are incomplete. Prior to inversion, three quality-control (QC) filters are automatically applied: (i) *clip*, which removes saturated traces; (ii) *ppsdoffline*, which identifies and excludes data exhibiting anomalous power spectral density (PSD) patterns relative to precomputed station envelopes; and (iii) *mouse*, which detects long-period disturbances or transient spikes and down-weights the affected time windows. These procedures ensure spectral consistency and inversion stability while maintaining low computational overhead, making them suitable for continuous real-time processing**.**

The *GISOLA* software—adapted from the *ISOLA* methodology (Zahradník, 2018)—performs the core moment tensor (MT) inversion using time-domain processing optimized for parallel execution on multi-core CPUs and GPUs. Inversion parameters such as window length, frequency range, and grid spacing are dynamically adjusted according to event magnitude and depth to achieve optimal signal representation. Time windows and frequency grids are synchronized with Fast Fourier Transform (FFT) parameters to ensure deterministic sampling and reproducible variance-reduction estimates. The inversion integrates waveform selection, noise weighting, and synthetic waveform matching within a four-dimensional grid-search space (latitude–longitude–depth–time), ensuring robust convergence even under complex network geometries.

Green's functions are generated using an Axitra-lineage engine executing concurrent CPU tasks. The model space is defined in both distance and depth, scaled according to magnitude—for example, M 3.0–4.0 events are computed using a 0–6 km distance and 0–21 km depth grid with 2 km spacing. Multiple one-dimensional velocity models are managed through a *Geobox* routing system. For Albanian events, the preferred model is *dushi_havskov.vz* (Dushi and Havskov, 2023), developed for the Albanian crustal structure, while *novotny.vz* serves as a fallback for events outside this region. This adaptive routing preserves

physical consistency across the Adria–Albanian transition, where lateral velocity variations are pronounced.

To maintain geometric stability, an azimuthal-coverage filter divides the 360° circle into eight equal sectors, requiring at least four sectors to be populated, with no more than two stations per sector after sorting by priority and data completeness. This constraint minimizes azimuthal gaps and mitigates mirrored or unstable mechanisms, particularly for smaller-magnitude events or asymmetric network configurations.

The main configuration parameters used in the automatic inversion workflow are summarized in Table 2, illustrating how distance, depth, frequency, and geometry criteria scale with magnitude in the Albanian implementation of the GISOLA system.

Table 2. Configuration parameters used in the automatic GISOLA inversion chain for near real-time MT computation within the Albanian Broadband Seismic Network. BB = Broadband; SM = Strong-motion channels.

Magnitude Range (Mw)	Distance Range (km)	Depth Grid (km)	Frequency Band (Hz)	Minimum No. of Stations	Azimuthal Sectors Required
3.0 – 4.0	0 – 350 (BB/SM)	0 – 21, step 2	0.05 – 0.20	≥ 6	≥ 4 of 8
4.0 – 5.0	0 – 800 (BB/SM)	0 – 31, step 2	0.03 – 0.15	≥ 8	≥ 4 of 8
5.1 – 6.0	up to 1000 (BB), 600 (SM)	0 – 31, step 2	0.02 – 0.12	≥ 10	≥ 4 of 8
6.1 – 7.0	up to 2000 (BB/SM)	0 – 31, step 2	0.02 – 0.10	≥ 12	≥ 4 of 8

Following inversion, GISOLA ranks candidate solutions according to correlation and variance-reduction scores. Each solution is automatically graded (A1–D4) based on waveform fit, spatial coverage, and numerical stability. Solutions achieving a grade of B4 or better are submitted back to the SeisComP database in SeisComP Markup Language (SCML) and QuakeML formats, disseminated through NCSM's web services, and shared in near real time with the Euro-Mediterranean Seismological Centre (EMSC, 2023) for integration into regional moment tensor (MT) catalogs. Published outputs include centroid coordinates, scalar moment (Mw), double-couple and compensated linear vector dipole (CLVD) components, and the orientation of the principal stress axes (T, P, and B).

The workflow follows a two-level revision policy. *Quick revisions* recompute only the inversion component under modified frequency bands, station configurations, or time windows, reusing previously computed Green's functions to maintain minute-scale latency. *Full reprocessing* is automatically triggered when upstream parameters—such as relocated origins, updated metadata, or revised crustal models—are introduced. All processing cycles preserve complete provenance information, including configuration files, inventory versions, and Green's-function identifiers, ensuring a transparent and auditable workflow for both operational monitoring and scientific validation.

From an implementation standpoint, the system operates on a virtualized, modular infrastructure designed for fault tolerance and scalability. *Core SeisComP* services (database, message broker, and SDS archive), real-time *SeedLink* streams, and *GISOLA* inversion engines function as distinct yet synchronized components within a secure local network environment. Internal communication is standardized via message passing and synchronized data directories, enabling multiple inversion nodes to execute in parallel during periods of increased waveform density. Configuration parameters are centrally versioned, while system performance is continuously tracked through diagnostic dashboards displaying latency, event throughput, and inversion success rates. Automatic notifications alert operators to anomalies such as unexpected delays, reduced waveform correlation, or degraded solution quality. These monitoring capabilities ensure the operational reliability of the national MT system and facilitate immediate corrective action when necessary.

This integrated architecture transforms the Albanian Broadband Seismic Network (ABSN) into a production-grade, high-performance system capable of delivering reliable Moment Tensor (MT) solutions within minutes of event detection. By combining SeisComP's rapid event-processing capabilities with GISOLA's efficient inversion engine and rigorous data-quality controls, Albania has established a reproducible, transparent, and scientifically validated framework for near real-time analysis of regional seismic sources. Beyond enhancing operational monitoring, this system supports seismotectonic interpretation and stress-field analysis across the Adria–Albanian orogenic contact. It also enables automated data dissemination: once the predefined quality threshold (≥ B4) is met, MT solutions are automatically transmitted to the EMSC for public catalog publication, uploaded to the IGEO web portal, and

announced via automated email notifications, typically within minutes of the seismic event.

4. Validation and Quality Assessment of the Moment Tensor Solutions

Validation of the near real-time Moment Tensor (MT) solutions generated by the GISOLA framework is conducted through a structured, multi-stage procedure that assesses both the internal consistency and seismotectonic reliability of each computed solution. The validation process integrates quantitative performance metrics derived from the inversion with independent fault-type classification using the Focal Mechanism Classifier (FMC) software (Álvarez-Gómez, 2019).

FMC is a command-line analytical tool designed to classify, cluster, and visualize double-couple (DC) components obtained from MT inversions. It computes derived parameters such as the principal stress axes (T, B, and P) and Kaverina coordinates (Kaverina, 1996), which project focal mechanisms into a barycentric diagram distinguishing normal, reverse, and strike-slip faulting regimes. Each event is assigned to a rupture type (thrust, normal, strike-slip, or oblique) and subsequently grouped using hierarchical clustering in both mechanism and geographic space. The resulting outputs include standardized classification diagrams, rose plots of strike–dip–rake parameters, and spatial density maps summarizing regional faulting patterns.

To ensure full compatibility with FMC, GISOLA maintains an event-based directory structure, in which each inversion result is archived together with its corresponding data products in multiple standardized formats (text, XML, QuakeML, and HTML). A custom parser developed at the Institute of Geosciences (IGEO) automatically converts all near real-time and revised MT solutions into FMC-compatible formats (Aki–Richards, CMT, and DC tensor representations). This integration enables seamless clustering, classification, and mapping directly from the operational repository, effectively transforming near real-time MT products into a standardized, research-grade dataset suitable for tectonic and seismotectonic interpretation.

The performance of the system was evaluated using the cumulative Moment Tensor (MT) catalogue generated since full operational deployment. These datasets not only reflect the current capabilities of the Albanian Broadband Seismic Network but also provide a benchmark for

assessing the reliability, stability, and internal consistency of the automated inversion workflow. Two main catalogues were analyzed: the automatic GISOLA catalogue containing *N = 1360* events, and the revised catalogue (*N = 137*), derived from targeted re-inversions and detailed quality-control procedures. A unified "best-per-event" catalogue was subsequently compiled by pairing identical timestamps and event identifiers, selecting the most reliable solution based on a hierarchical criterion that prioritizes (i) quality class ($A < B < C < D$), (ii) waveform correlation, and (iii) variance reduction (VR). The resulting merged dataset, comprising *N = 1360 events*, constitutes the most thoroughly validated MT archive currently available for the Albanian region (Fig. 7).

Magnitude ranges confirm the robustness of the inversion and consistency of the revision workflow:

- Automatic catalogue : Mw 2.9–6.4 (Origin magnitude : 2.5–6.4)
- Revised catalogue: Mw 2.9–6.2 (Origin magnitude: 2.7–6.4)
- Merged catalogue : Mw 2.9–6.4 (Origin magnitude : 2.5–6.4)

A total of 88 events appear in both catalogues. Within this overlap, revised solutions are preferred in *66 out of 88* cases (≈ 75%), confirming that the revision process systematically improves waveform fitting and inversion stability. Quantitative comparisons reveal consistent enhancement across key metrics:

- Δ Quality-class rank = +10.0 (positive = improvement)
- Δ Correlation = +0.122
- Δ Variance reduction = +0.148
- Median centroid shift = 1.31 km

These indicators demonstrate significant improvement in waveform correlation and variance reduction, with negligible magnitude bias and minor centroid shifts, the latter reflecting geometric refinement rather than inversion instability. Catalogue-wide statistics further support this trend:

- Automatic (N = 1360): median correlation ≈ 0.32; VR ≈ 0.10; A-class ≈ 6–7 %.
- Revised (N = 137): median correlation = 0.705; VR = 0.484; A-class ≈ 24 %.
- Merged (N = 1360): median correlation ≈ 0.32; VR ≈ 0.10; A-class ≈ 7 %.

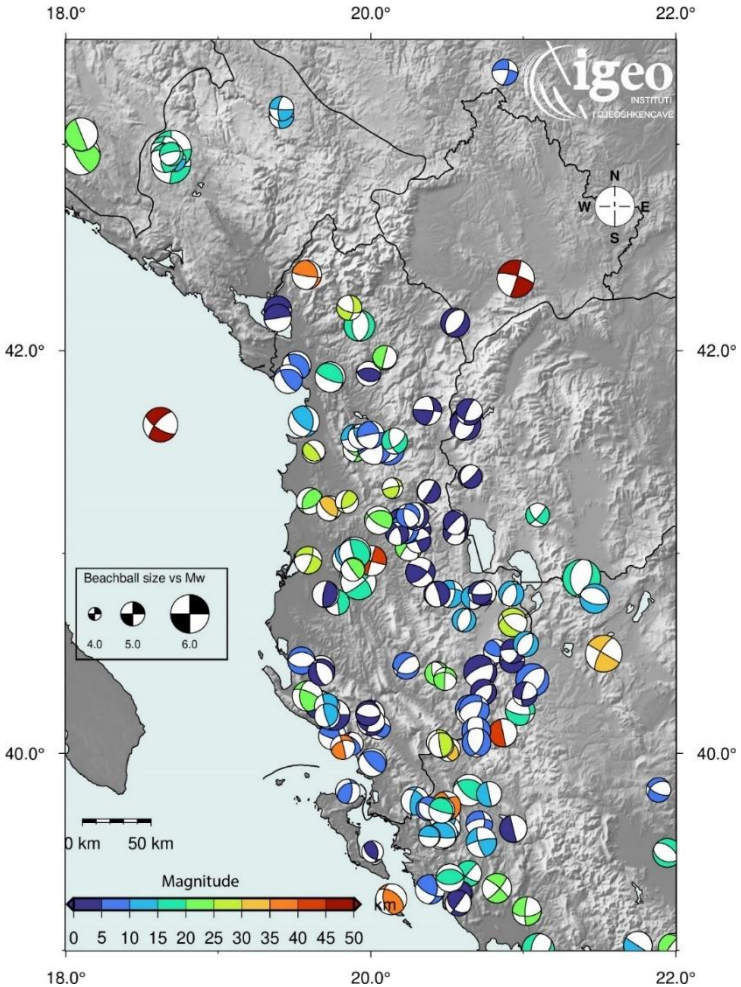


Fig. 7. GISOLA moment tensor solutions for 248 events with quality B4 and better, within the polygon of the near-Albania region (latitude 39–43.5° N, longitude 18–22° E).

The distribution of quality grades across all events (Fig. 8) exhibits a pronounced shift toward higher-confidence classes following the revision process, confirming that systematic quality-control cycles substantially improve the reliability of near real-time MT determinations.

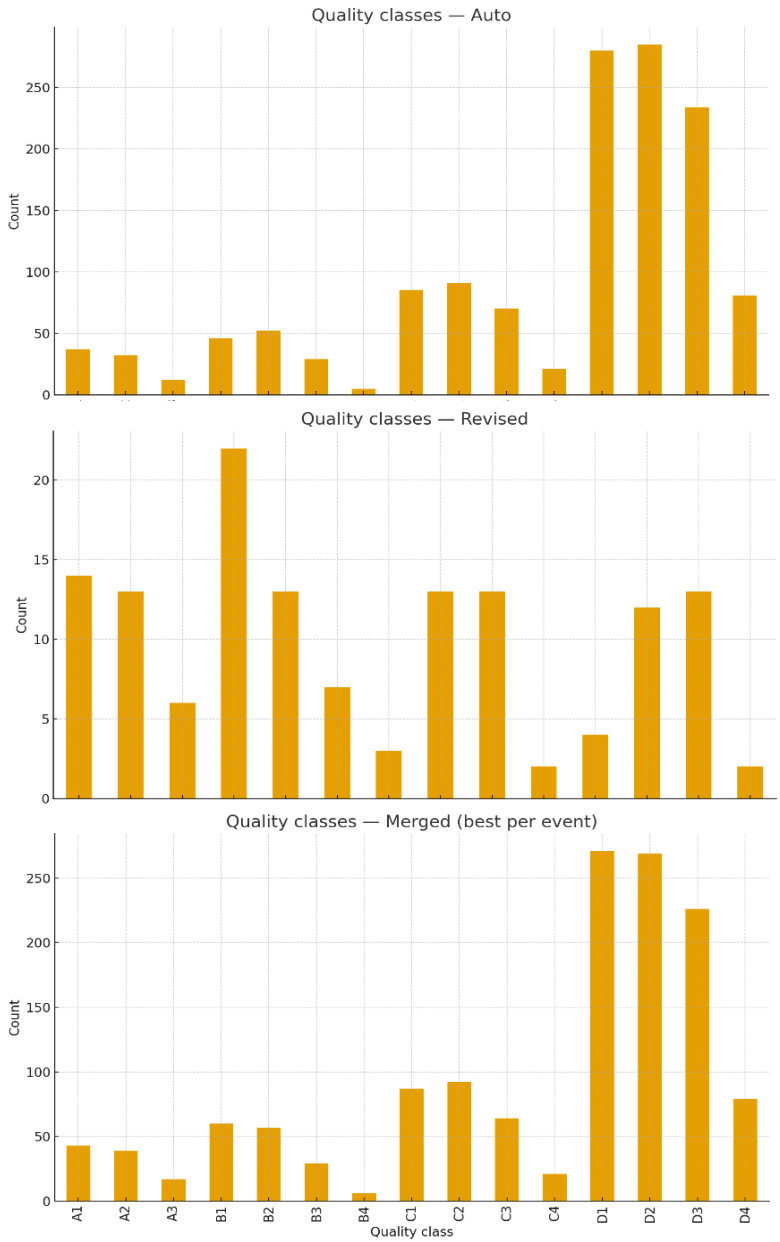


Fig. 8. Number of MT solutions by quality bin, from A1 (best) to D4 (worst). The revised dataset shows a marked increase in higher-quality (A–B) solutions relative to the automatic catalogue.

To further interpret the seismotectonic framework, *FMC* classification was applied to all MT solutions of grade B4 or better within the region 39–43.5° N and 18–22° E. *FMC* classifies focal mechanisms by rupture type and clusters them according to similarity within the Kaverina coordinate system. The resulting Kaverina diagram (Fig. 9) delineates three principal rupture clusters—Normal (N), Reverse (R), and Strike-Slip (SS)—with transitional oblique categories (N–SS, N–R, R–SS) bridging them. Reverse faulting mechanisms predominate along the Western Albanides, whereas strike-slip and oblique mechanisms become more frequent toward the internal orogenic belt, consistent with the present-day compressional to transpressional regime that characterizes the Adria–Albanian orogenic contact.

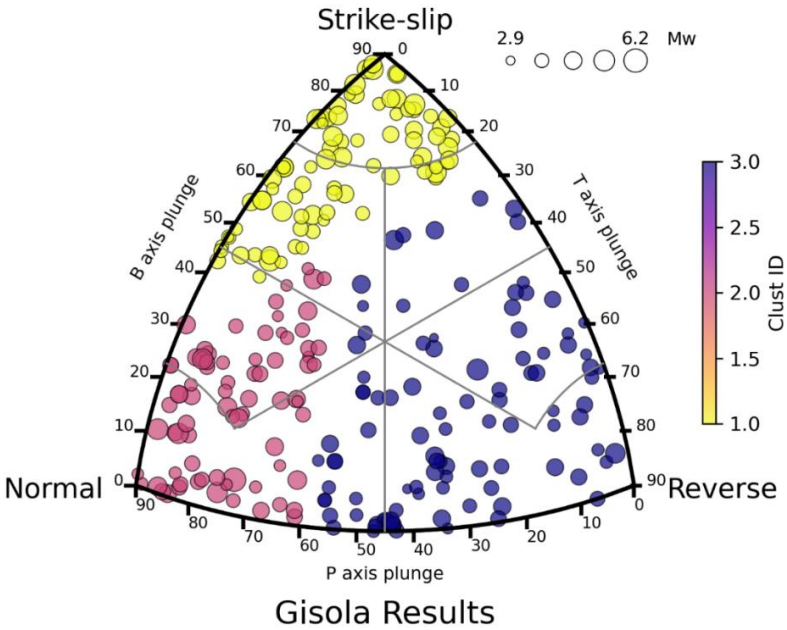


Fig. 9. Kaverina plot for all GISOLA-derived moment tensor solutions within the near-Albania region. Three principal clusters are identified, corresponding to dominant Normal, Reverse, and Strike-Slip regimes (Kaverina *et al.* 1996).

5. Discussion

5.1 Operational and methodological improvements

The implementation of the near real-time GISOLA–FMC framework represents a major advancement in Albania's seismic monitoring

capability. Quantitative analyses indicate that azimuthal-gap reductions of 10–25° across events in the magnitude range M 0.5–4.5, combined with an improved magnitude of completeness (Mc ≈ 1.35 in 2024), have extended the domain of stable Moment Tensor (MT) inversions to smaller-magnitude events, historically constrained by network geometry. The introduction of a two-tier revision cycle and an automated quality-class system has further enhanced result consistency, yielding median improvements of +0.122 in waveform correlation and +0.148 in variance reduction across overlapping events, while maintaining Mw stability and minimal centroid shifts (~1 km). These metrics reflect a calibrated refinement process rather than wholesale re-parameterization, ensuring both reproducibility and traceability of solutions.

The resulting "best-per-event" catalogue constitutes a coherent, high-quality dataset suitable for operational applications, including ShakeMaps, fault-plane solutions, and near real-time stress-field monitoring. The integration of automated inversion, systematic quality control, and mechanism classification has transformed the Albanian Broadband Seismic Network into a robust, transparent, and scientifically auditable system, comparable to advanced regional frameworks across Europe.

5.2 Tectonic and neotectonic validation

The tectonic framework of Albania is controlled by the ongoing convergence between the Adriatic (Adria) and Eurasian plates, producing a continuum of deformation from the external compressional front to the internal extensional basins. The neotectonic architecture, as defined by Aliaj (1996), is dominated by major fault systems, including the Ionian thrust front, and the Shkodra–Lezha, Tirana–Elbasan, and Drini fault zones, which have accommodated differential motion between crustal blocks since the Pliocene–Quaternary. The spatial distribution of the most reliable GISOLA MT solutions aligns closely with these active structures, confirming both the robust performance of the automated system and the geological realism of its results.

Along the western margin, the predominance of reverse and oblique-thrust mechanisms corresponds to the Ionian fold-and-thrust belt, a zone of active crustal shortening consistent with the Pliocene–Recent compressional front mapped by Aliaj (1996) and the geodetic shortening rates of 3–5 mm yr^{-1} reported by Jouanne et al. (2012). This spatial and kinematic concordance confirms that the near real-time MT solutions

faithfully reproduce the active deformation of the External Albanides, validating the operational workflow against independent geological and geodetic constraints.

Toward central Albania, the strike-slip and transpressional mechanisms observed along the Tirana–Elbasan–Dibra zone delineate a major transfer corridor linking the compressional front with the internal extensional basins. The predominance of right-lateral shear along this corridor corroborates the transition from contractional to transtensional kinematics inferred by Jouanne (2012) and earlier seismotectonic interpretations by Aliaj (1996).

In northeastern Albania, the normal and oblique-normal faulting solutions clustered along the Drini Fault Zone substantiate its role as an active extensional boundary. The MT-derived fault geometries (strike ≈ NNW–SSE; dip 50–60°) align with mapped neotectonic scarps and with moderate-magnitude seismicity (Mw 3–4) recorded near Peshkopi and Kukës. This correspondence reinforces the long-term extensional regime described by Aliaj (1996) and demonstrates that the GISOLA–FMC solutions accurately capture genuine crustal deformation within the Internal Albanides, rather than representing artefacts of the inversion process.

5.3 Implications for regional seismotectonics

The spatial distribution of validated MT mechanisms (Fig. 10) delineates three major structural domains that characterize the active deformation of Albania:

1. **Western compressional domain (Ionian–Durrës–Vlora belt):** dominated by reverse and thrust faulting within the external fold-and-thrust system.
2. **Central transpressional corridor (Tirana–Elbasan–Dibra zone):** defined by strike-slip and oblique mechanisms, marking the transitional boundary between compression and extension.
3. **Eastern extensional domain (Korça–Ohrid–Drini region):** characterized by normal faulting**,** consistent with crustal thinning and Pliocene–Quaternary basin formation**.**

This tectonic segmentation closely reflects both the neotectonic zonation established by Aliaj (1996) and the strain-partitioning patterns quantified by Jouanne (2012). The strong correspondence between the automatic MT classifications and these independent datasets validates the

GISOLA–FMC operational framework as a scientifically robust system capable of resolving first-order deformation processes across the Adria–Albanian orogenic contact**.**

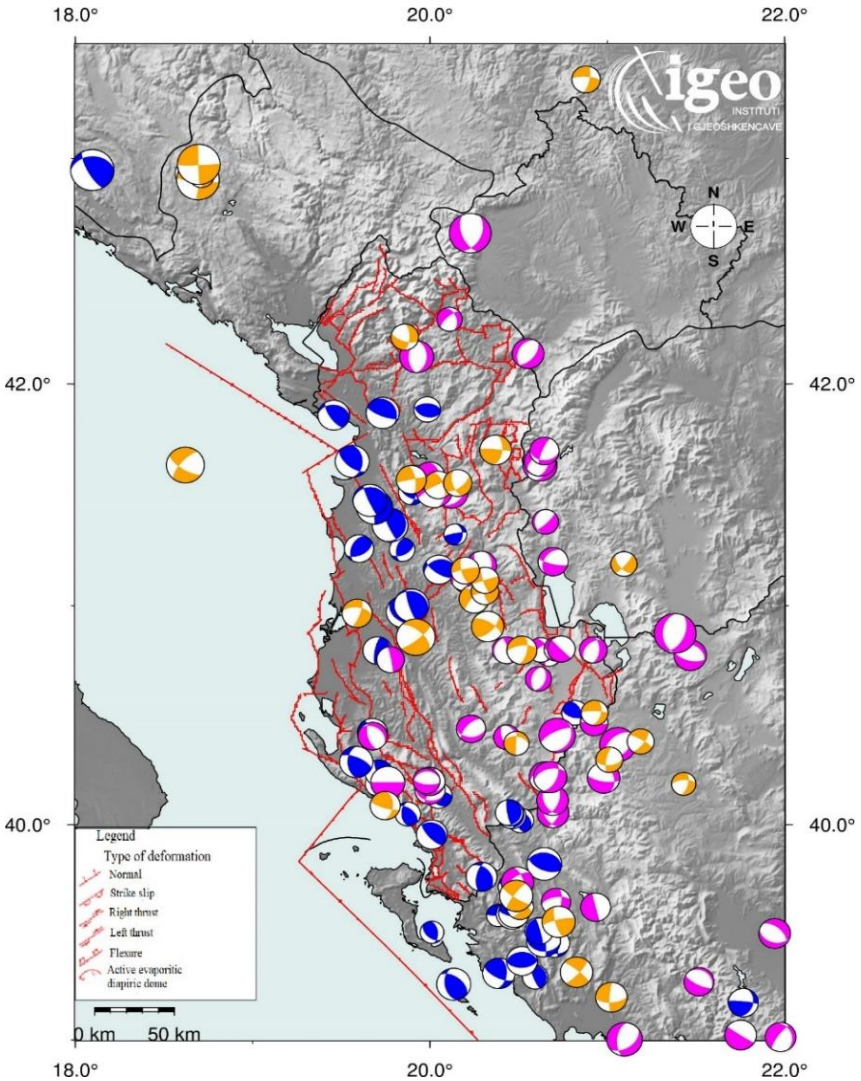


Fig. 11. Moment Tensor (MT) solutions clustered with FMC. Reverse mechanisms are shown in blue, normal in magenta, and strike-slip in orange, illustrating the spatial distribution of rupture styles and the prevailing tectonic regime across Albania.

The improved network geometry, lower completeness magnitude, and systematic quality assessment collectively ensure that the near–real-time MT solutions provide a robust and physically meaningful record of ongoing crustal deformation across Albania. Consequently, the national system not only enhances rapid-response capabilities but also establishes a dependable foundation for long-term seismotectonic and geodynamic research, directly supporting the regional objectives of the AdriaArray initiative**.**

6. CONCLUSIONS

This study establishes the first production-grade, near-real-time Moment Tensor (MT) framework for Albania, integrating *SeisComP* for event logistics, *GISOLA/ISOLA* for automated inversion and revision, and *FMC* for classification and clustering. The system combines robust procedures—automatic data selection, QC filtering, and structured revision—with standardized products such as an auditable MT catalogue, QuakeML exports, and tectonic mechanism families. It represents a major transition from ad hoc analyses to continuous, uncertainty-aware MT operations adapted to 24/7 monitoring.

Network densification following the 2019 Durrës earthquake reduced azimuthal gaps by 10–25° and lowered the completeness magnitude (Mc) from ~2.25 to ~1.35, extending the domain of stable inversions to events as small as M 2.5–3.0. These improvements have directly enhanced MT stability, reproducibility, and waveform-fit quality. The two-tier revision scheme further improved correlation and variance reduction without introducing magnitude bias, confirming that the process yields targeted and quantifiable refinements rather than reparameterization.

The curated MT catalogue now provides a reliable, reproducible, and spatially coherent record of active deformation, accurately delineating Albania's principal tectonic domains—compressional faulting along the Ionian front, transpressional deformation along the Tirana–Elbasan corridor, and normal faulting within the Drini and Korça–Ohrid basins**.** The framework's integration with *EIDA/ORFEUS* ensures data accessibility, interoperability, and scalability within the broader AdriaArray initiative**.**

This workflow establishes a novel operational model for near–real-time seismic source characterization across the Adria–Albanian orogenic contact**.** By bridging the gap between seismic monitoring and tectonic

research, it delivers standardized, validated MT products that simultaneously support early warning, hazard assessment, and long-term geodynamic investigations. Future developments—including regionally optimized velocity models and cross-border catalogue harmonization—will further strengthen Albania's role as a regional contributor to real-time seismology in the central Mediterranean.

ACKNOWLEDGMENTS

The authors acknowledge the Institute of Geosciences (IGEO) for institutional support and the Albanian Seismic Network operations team for their dedicated efforts in maintaining the national broadband infrastructure. Network improvements were supported by the RESEAL Project (Resilience Strengthening in Albania). The authors also express gratitude to the anonymous reviewers for their constructive feedback and valuable suggestions that improved the quality of this work.

Declarations

Data Accessibility

All data used in this study are publicly available and have been previously published or referenced in scientific sources. No additional or proprietary datasets were generated or analyzed during the current study. Detailed dataset may be made available upon reasonable request to the authors.

Declaration of AI Use

AI was not used for idea conception, manuscript writing, data interpretation, or analytical reasoning. All scientific content, analyses, and conclusions were independently conceived and verified by the authors.

Authors' Contributions

Rama B-conceptualization, writing, original draft, data collection, figure development and reference validation;

Dushi E- seismicity and seismotectonics analysis, validation and reviewing;

Koxhaj D- verification from published and official open sources.

Author Approval

All authors have read and approved the final version of the manuscript and agree to be held accountable for all aspects of the work.

Conflict of Interest Declaration

The authors declare no conflict of interest.

Funding

This research received no external funding.

REFERENCES

Aliaj SH. 2020. Seismotectonics of the albanides collision zone: geometry of the underthrusting adria microplate beneath the albanides. *Albanian Journal of Natural and Technical Sciences* (AJNTS) **2 (51):** 3-41. https://akad.gov.al/albanian-journal-of-natural-and-technical-sciences-ajnts-20202-vol-xxv-51/.

Aliaj Sh. 2006. The Albanian orogen: convergence zone between Eurasia and the Adria microplate." *Springer Netherlands* (Springer Netherlands) 133–149. DOI https://doi.org/10.1007/1-4020-4235-3_09.

Aliaj Sh, Kociu S, Muco B, Sulstarova E. 2010. Seismicity, seismotectonics and seismic hazard assessment in Albania. *The Academy of Sciences of Albania* (The Academy of Sciences of Albania). ISBN: 978-99956-10-26-5.

Álvarez-Gómez, José A. 2019. FMC—Earthquake focal mechanisms data management, cluster and classification. *SoftwareX* **9(93):2**99-307 DOI:10.1016/j.softx.2019.03.008.

Dushi E, Rama B, Schlömer A, Egdorf E,Terpoorten M. 2025. AdriaArray in Albania: Collaborative deployment, station characterization, and data integration. *Annals Of Geophysics,* **65**:1-5.https://doi.org/10.4401/ag-9303.

Dushi E, Havskov J. 2023. 1D c rustal structure of Albania region. *Annals of Geophysics,* **66 (1):** SE103. https://doi.org/10.4401/ag-8805.

Euro-Mediterranean Seismological Centre. 2023. *EMSC.* https://www.emsc-csem.org/.

Evangelidis CP, Triantafyllis N, Samios M, Boukouras K, Kontakos K, Ktenidou O-J, Fountoulakis I, Kalogeras I, Melis NS, Galanis, O, Papazachos CB, Hatzidimitriou P, Skordilis M, Sokos E, Paraskevopoulos P, Serpetsidaki. 2021. Seismic waveform data

from Greece and Cyprus: Integration, archival and open access. *Seismological Research Letters,* **92(3):** 1672–1684. doi: 10.1785/0220200408.

Helmholtz-Centre Potsdam - GFZ German Research Centre for Geosciences and gempa GmbH. 2008. *The SeisComP seismological software package. GFZ Data Services.* doi:10.5880/GFZ.2.4.2020.003.

Institute of Geosciences. 2023. *Seismological Network at IGEO.* https://geo.edu.al/Research/Research_Infrastructure/Seismology_Network/.

Institute of GeoSciences, Polytechnic University of Tirana. 2002. *Albanian Seismological Network.* https://www.fdsn.org/networks/detail/AC/.

Jouanne F, Mugnier J-L, Koci R, Bushati S, Matev K, Kuka N, Shinko I, Kociu S, D Ll. 2012. GPS constraints on current tectonics of Albania. *Tectonophysics*, **554–557: 50-62 https://doi.org/10.1016/j.tecto.2012.06.008.**

Kaverina AN, Lander AV, Prozorov AG. 1996. Global Creepex distribution and its relation to earthquake-source geometry and tectonic origin. *Geophysical Journal International*, **125 (1):** 249–265, https://doi.org/10.1111/j.1365-246X.1996.tb06549.x.

Moshou A, Dushi E, Argyrakis P. 2019. *A preliminary report on the 26 november 2019, Mw=6.4.* emsc-csem.org.

Strollo A, Cambaz D, Clinton J, Danecek P, Evangelidis ChP, Marmureanu A, Ottemöller L, Pedersen H, Sleeman R, Stammler K, Armbruster D, Bienkowski J, Boukouras, K., Evans, P., Fares M, Neagoe C, Heimers S, Heinloo A, Hoffmann M, Kaestli P, Lauciani V, Michalek J, Odon Muhire E, Ozer M, Palangeanu L, Pardo C, Quinteros J, Quintiliani M, Antonio Jara-Salvador J, Schaeffer J, Schloemer A, Triantafyllis N. 2021. EIDA: The European Integrated Data Archive and Service Infrastructure within ORFEUS. *Seismological Research Letters,* **92 (3):** 1788–1795 https://doi.org/10.1785/0220200413.

Suarez G, van Eck T, Giardini D, Ahern T, Butler R, Tsuboi S. 2008. The International Federation of Digital Seismograph Networks (FDSN): An Integrated System of Seismological Observatories. *IEEE Systems Journal*, **2(3):** 431-438. DOI 10.1109/JSYST.2008.2003294.

Teloni S, Invernizzi Ch, Mazzoli S, Pierantoni PP, Spina V. 20 The seismogenic fault system of the M w 6.4, November 2019 Albania

earthquake: new insights into the structural architecture and active tectonic setting of the outer Albanides *Journal of the Geological Society*, **178(2):**jgs2020-193 DOI:10.1144/jgs2020-193.

Triantafyllis N, Venetis IE, Fountoulakis I, Pikoulis E, Sokos E, Evangelidis CP. 2021. Gisola: A High-Performance Computing Application for Real-Time Moment Tensor Inversion. *Seismological Research Letters*, 93 (2A): 957–966. https://doi.org/10.1785/0220210153957–966.

Vittori E, Blumetti AM, Comerci V, Di Manna P, Piccardi L, Gega D, Hoxha I. 2020. Geological effects and tectonic environment of the 26 November 2019, Mw 6.4 Durres earthquake (Albania). *Geophysical Journal International*, **225 (2):**1174-1191. 10.1093/gji/ggaa582.

Wessel P, Luis JF, Uieda L, Scharroo R, Wobbe F, Smith WHF, Tian D. 2019. The generic mapping tools Version 6. *Geochemistry, Geophysics, Geosystems,* **20(11):** 5556–5564. https://doi.org/10.1029/2019GC008515.

Zahradník J, Sokos E. 2018. ISOLA Code for Multiple-Point Source Modeling—Review. *Springer Natural Hazards* 1-28. 10.1007/978-3-319-77359-9_1.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

COSEISMIC STRESS IN THE 2019 DURRËS SEISMIC SEQUENCE: TENSOR RESULTS AND EARLY AFTERSHOCK ADJUSTMENT

Edmond DUSHI, Besian RAMA, Damiano KOXHAJ
Department of Seismology, Institute of Geosciences (IGEO), Polytechnic University of Tirana, Albania
Author for correspondence e.dushi@geo.edu.al
ED, 0000-0001-7599-7034; BR, 0000-0002-6759-2130;
DK, 0009-0002-9792-9010
https://doi.org/10.67537/ajnts3001003

ABSTRACT

We quantify the coseismic stress field of the 26 November 2019 Durrës earthquake by integrating a robust solution obtained with TENSOR for the mainshock and an early post-seismic data inversion with the StressInverse software. The inversion was applied to the highest-magnitude events in the early sequence (36 mechanisms reported in the Euro-Mediterranean Seismological Centre Preliminary Report issued immediately after the earthquake). A regional inversion based on published focal mechanisms and moment tensor solutions for the earthquakes of 1964–2019 defines the broader tectonic framework. Across these datasets, we estimate the orientations of the principal stress axes (σ_1, σ_2, σ_3), together with the stress shape ratio (φ) and apparent friction (μ), which describe the stress hierarchy and internal fault friction governing deformation. The mainshock exhibits reverse faulting under WSW–ENE compression, with $\sigma_1 \approx$ 249/23, $\sigma_2 \approx$ 157/5, $\sigma_3 \approx$ 56/66 (azimuth/plunge, degrees), $R \approx 0.5$, and quality class “B”. The early sequence preserves the compression azimuth but shifts to $\varphi \approx 0.55$ and $\mu \approx 0.45$, consistent with coseismic weakening and transient stress redistribution. The regional field remains dominantly compressional, while sequence-time inversions reveal a short-lived perturbation induced by the Mw 6.3 rupture. Overall, the results indicate permanent crustal shortening accommodated along thrust and back-thrust faults due to the ongoing frontal continental subduction of the Adria microplate confronting the Albanian orogeny, and demonstrate that integrating immediate aftershock mechanisms enhances the temporal resolution of stress evolution and fault mechanics in active collision zones.

***Keywords*:** Durrës 2019, coseismic stress; TENSOR, StressInverse, aftershocks, reverse faulting

1. INTRODUCTION

The Durrës region lies within the external (Outer) Albanides, the northward continuation of the Hellenides fold-and-thrust belt, at the active

margin of the Adria microplate confronting the Eurasian domain. This foreland belt accommodates NW–SE shortening (i.e., WSW–ENE principal compression) driven by Adria–Eurasia convergence, expressed through regionally continuous thrusting and associated folding (Meço and Aliaj, 2000; D'Agostino *et al*. 2008; Jouanne *et al*., 2012; Métois *et al*. 2015; Matraku *et al*. 2023). A seismotectonic framework for the study area is presented in Fig. 1, with the inset (b) illustrating the Adria–Eurasia collision boundary from Aliaj (2016).

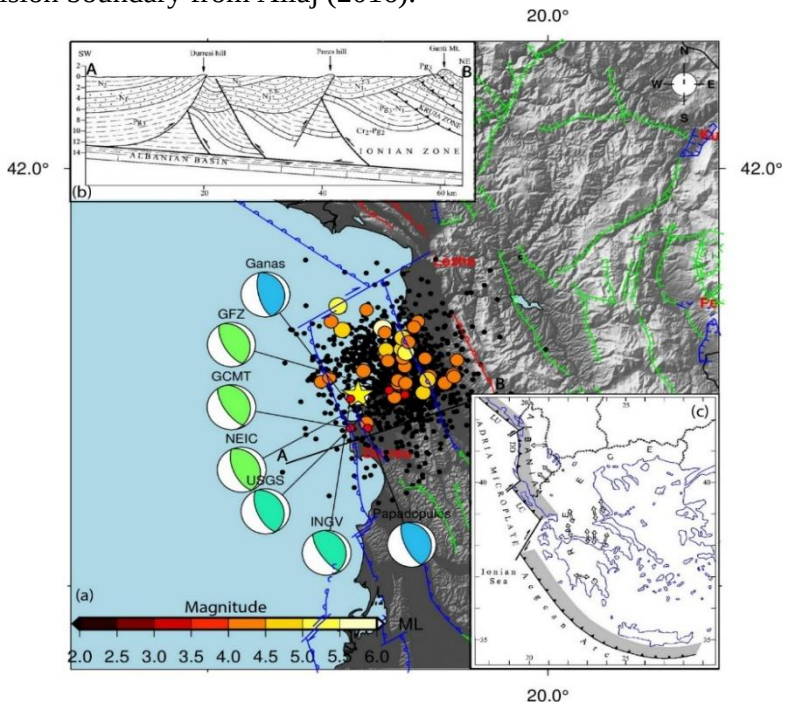


Fig. 1. (a) Map of the 26 November 2019 Durrës earthquake sequence (Mw 6.4 mainshock) with aftershocks and representative focal mechanisms used in the stress inversions, compiled from the Institute of Geosciences (IGEO) internal catalogue and Euro-Mediterranean Seismological Centre (EMSC) moment-tensor database (https://www.emsc-csem.org/Earthquake_data/tensors.php). (b) Generalized schematic of the southern convergent margin of the Eurasia–Adria collision zone, modified after Aliaj (2016), showing the segmentation of the Albanides collision front into LC (Lefkas–Corfu) and DD (Dhërmi–Durrës); LU (Lezhë–Ulqini) corresponds to the Dinarides s.str. (c) Geological cross-section through Durrës Hill–Prezë Hills–Ganti Mt., illustrating head-on back thrusting within the frontal Albanides.

Persistent seismicity along the Adria–Albanides collision margin culminated on 26 November 2019 with the Durrës mainshock (Mw 6.4), followed by a dense aftershock sequence of more than 1000 events delineating the activated volume (Ganas *et al*., 2020; Papadopoulos *et al*., 2020; Mavroulis *et al*., 2021). This seismic crisis provides a natural context for contrasting the ambient (regional) stress field with the transient conditions that accompany and immediately follow a major rupture. In this work, we concentrate on the coseismic state associated with the mainshock and on the very-early post-seismic adjustment captured by the mainshock and 36 immediate aftershocks (Moshou and Dushi, 2019; EMSC, 2024). Routine monitoring and joint analyses by national and regional institutions, particularly the Institute of Geosciences, Polytechnic University of Tirana (IGEO) and the National Observatory of Athens (NOA), provide the focal-mechanism and moment-tensor solutions used in this work, ensuring compatibility and cross-validation across catalogs (Konstantinou *et al.* 2010; Roumelioti *et al*. 2011; Triantafyllis *et al*. 2022). For the mainshock, we compile a multi-agency focal-plane ensemble (NOA, IGEO, GFZ, GCMT, INGV, AUTH, and USGS) (Table 1). For the early sequence, we adopt the mainshock and 36 aftershocks from the immediate post-event evaluation by Moshou and Dushi (2019; EMSC), which provides the mechanisms used in our sequence analysis (Table 4). The instrumental context is summarized in Figure 2, which shows the current broadband seismic network (IGEO Seismological Network; http://www.orfeus-eu.org/data/eida/) used for regional monitoring, with an inset illustrating the sparser configuration operating at the time of the 2019 earthquake, differences that bear directly on mechanism availability and quality.

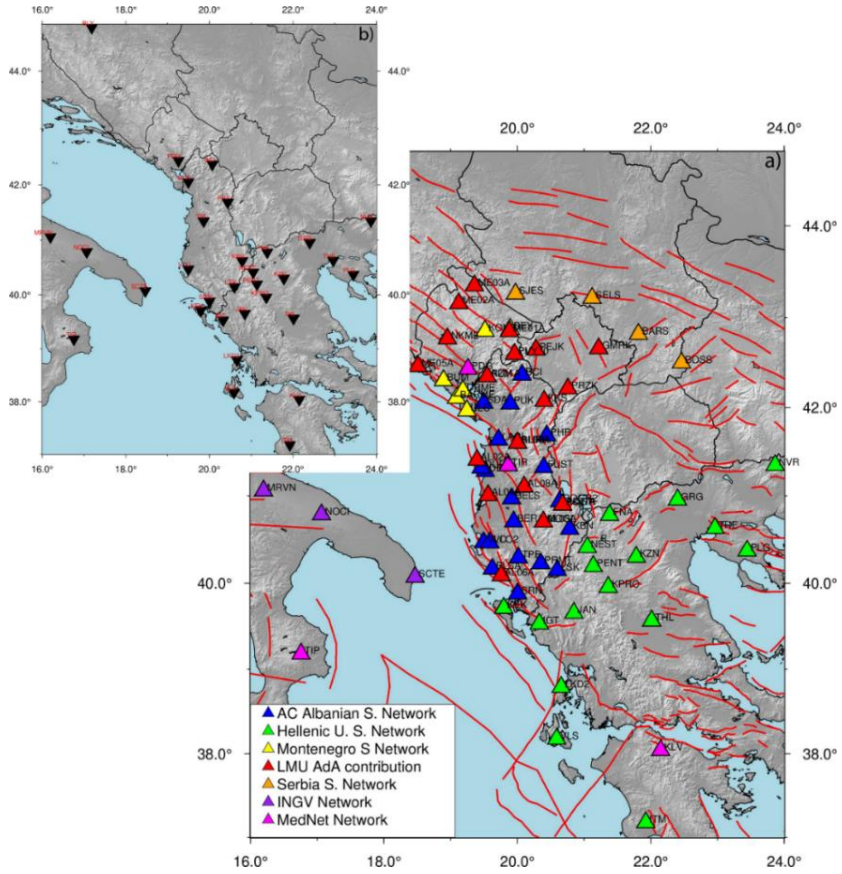


Fig. 2. (b) Distribution of permanent broadband seismic stations operated by IGEO and contributing to the EIDA (ORFEUS) archive, defining the present Albanian Seismological Network. (a) Historical configuration of the network active during the 2019 Durrës earthquake. Scale bar and north arrow are shown for reference. Data from IGEO, NOA, and EIDA (ORFEUS).

Over recent decades, the Albanian stress field has been investigated using first-motion focal mechanisms and routine moment-tensor solutions compiled across agencies and catalogues (Muço, 1994; Muço, 2007; Kiratzi and Dimakis, 2013; Dushi *et al*. 2018). Against this background, Teloni *et al*. (2021) delineated the seismogenic fault system responsible for the 2019 sequence, documenting a reverse-fault-dominated architecture consistent with Outer Albanides compression, and emphasizing the structural segmentation active during the crisis. Their

structural interpretation provides a benchmark for the present stress-based analysis, which is further contextualized by regional GPS-constrained deformation patterns (Jouanne *et al*. 2012; Métois *et al*. 2015; D'Agostino *et al*. 2008) and by the coseismic and post-seismic deformation field derived from geodetic and InSAR data (Matraku *et al*. 2023).

In this study, we use only published mechanisms, and no new focal solutions are computed, to characterize (i) the coseismic stress tensor of the 2019 mainshock and (ii) the very-early post-seismic stress state derived from the mainshock and 36 aftershocks. We further compare these with a regional, long-term inversion constructed from published sources (a proxy for the permanent regional stress field) and with previous findings, notably those by Teloni *et al*. (2021). Emphasis is placed on stress orientation (principal axes σ_1, σ_2, σ_3) and strength proxies—the stress shape (φ) and apparent friction (μ) — to evaluate how the Durrës source zone evolved through the onset of the sequence. This comparative approach, involving stress inversion of representative moment-tensor and focal-mechanism datasets from distinct time windows relative to the Durrës earthquake, aims to reveal local stress-field modifications during the seismic activity and to provide a robust baseline for future updates as monitoring and catalogues continue to improve.

2. DATA

This study resolves the stress field across three windows associated with the 2019 Durrës seismic crisis: (i) the coseismic state of the mainshock, (ii) the very-early post-seismic state defined by the mainshock plus 36 immediate aftershocks (26 November–2 December 2019), and (iii) a regional/tectonic state representative of the Durrës–Tirana seismotectonic environment. All inputs are published focal-mechanism (FM) and moment-tensor (MT) solutions treated as fixed; no waveform re-inversions were performed. The operational backbone is the NOA/IGEO catalogue, whose routine MTs are generated with ISOLA/GISOLA (Sokos and Zahradník, 2008; Konstantinou *et al*. 2010; Triantafyllis *et al*. 2022). However, the catalogue still reflects certain limitations related to azimuthal coverage and the variable sensitivity of seismic stations across the network.

The integration of data from multiple international and regional agencies (NOA, IGEO, GFZ, GCMT, INGV, AUTH, and USGS) ensures high reliability and cross-validation of the results. For the mainshock, we

assembled a multi-agency focal-plane ensemble from these institutions, retaining one nodal pair per entry, recording any preferred movement-plane flag, and screening for duplicates and parameter plausibility (origin time, magnitude, depth). These inputs feed the coseismic TENSOR analysis and are listed in Table 1; their spatial context and the resulting TENSOR axes are shown in Figure 1.

All waveform data are archived through the European Integrated Data Archive (EIDA; ORFEUS), part of the International Federation of Digital Seismograph Networks (FDSN), while the corresponding moment-tensor (MT) parameters are available from the Euro-Mediterranean Seismological Centre (EMSC).

Table 1. Mainshock (26 Nov 2019) focal-plane inputs used in TENSOR.

Agency	(°) Strike (degrees)	(°) Dip (degrees)	(°) Rake (degrees)	**Mw**
NOA (IGEO/NOA)	155	72	85	6.3
IGEWE / IGEO	339	27	111	6.4
GFZ	151	72	89	6.4
GCMT	349	24	111	6.4
INGV	352	22	115	6.4
AUTH	150	49	109	6.4
USGS	337	27	91	6.4

The 36 aftershocks used for the sequence inversion were obtained from the immediate post-event evaluation by Moshou and Dushi (2019, EMSC). In brief, moment tensors (MTs)—typically for events with ML ≥ 4.0— were computed using local and regional broadband waveforms retrieved from EIDA (ORFEUS). For regional inversions, a minimum of four stations with good azimuthal coverage and epicentral distance ≤ 350 km were required. Pre-processing included band-pass filtering (0.02–0.05 Hz), instrument response removal, integration to displacement, and rotation of horizontals to the back-azimuth. Green's functions were computed using the reflectivity method (Kennett, 1983) as implemented by Randall (1994) (workflow details in Kyriakopoulos *et al*. 2013).

Source depth was first scanned on a 5 km grid to locate the misfit minimum and then refined to 1–2 km around the best depth, using an appropriate 1-D velocity model (Drakatos *et al*. 2002). Solution quality was assessed based on the average misfit and CLVD percentage. The MTs

for the mainshock and the 36 aftershocks used in our StressInverse analysis are compiled in Table 2, and their distribution relative to the structural framework is displayed in Fig. 1.

Table 2. Mainshock and 36 immediate aftershocks (26 Nov–2 Dec 2019) used in StressInverse program (origin time, location, depth, nodal plane, Mw, reported Mo).

#	Date	Time	Lat (deg.)	Lon (deg.)	Depth (km)	Strike (deg.)	Dip (deg.)	Rake (deg.)	Mw	Mo (dyne·cm)
1	26/11/2019	02:54	41.459	19.442	25	155	72	85	6.3	5.20E+25
2	26/11/2019	02:59	41.504	19.652	18	145	65	80	4.4	5.12E+22
3	26/11/2019	03:03	41.509	19.722	17	143	72	84	4.4	5.12E+22
4	26/11/2019	03:21	41.566	19.615	20	151	70	86	3.9	9.80E+21
5	26/11/2019	03:57	41.612	19.498	26	146	73	80	3.9	9.80E+21
6	26/11/2019	04:22	41.528	19.579	27	157	66	89	4.2	2.72E+22
7	26/11/2019	05:35	41.580	19.469	23	140	68	85	3.8	5.74E+21
8	26/11/2019	05:50	41.558	19.683	19	146	50	95	3.8	5.74E+21
9	26/11/2019	05:56	41.549	19.574	21	140	72	88	4.4	5.12E+22
10	26/11/2019	06:08	41.619	19.514	24	133	64	67	5.5	2.61E+24
11	26/11/2019	06:20	41.579	19.687	25	155	63	89	3.9	9.80E+21
12	26/11/2019	07:12	41.609	19.522	26	158	60	88	4.0	1.48E+22
13	26/11/2019	07:27	41.569	19.569	24	166	62	95	4.6	8.80E+22
14	26/11/2019	07:36	41.454	19.549	19	143	71	94	4.2	2.72E+22
15	26/11/2019	12:14	41.484	19.581	19	155	72	87	4.0	1.48E+22
16	26/11/2019	13:05	41.596	19.573	15	159	75	89	4.9	2.70E+23
17	26/11/2019	17:06	41.566	19.587	21	143	65	90	4.1	1.84E+22
18	26/11/2019	17:19	41.615	19.396	25	154	65	85	4.9	2.70E+23
19	26/11/2019	19:44	41.599	19.616	22	159	76	83	3.8	5.74E+21
20	27/11/2019	11:03	41.552	19.600	29	146	73	91	3.9	9.80E+21
21	27/11/2019	14:45	41.672	19.383	23	155	56	86	5.3	1.18E+24
22	27/11/2019	22:19	41.635	19.363	23	160	75	85	4.2	2.72E+22
23	27/11/2019	22:50	41.517	19.458	21	149	73	85	4.4	5.12E+22
24	27/11/2019	22:51	41.487	19.559	17	146	52	97	4.2	2.72E+22
25	27/11/2019	23:02	41.518	19.411	16	159	32	90	3.8	5.74E+21

26	28/11/2019	00:50	41.663	19.467	15	167	65	96	4.0	1.48E+22
27	28/11/2019	10:25	41.552	19.524	25	145	67	92	4.1	1.84E+22
28	28/11/2019	10:52	41.561	19.584	25	159	31	89	5.0	3.75E+23
29	28/11/2019	13:28	41.635	19.618	21	145	73	90	4.0	1.48E+22
30	28/11/2019	14:22	41.491	19.333	26	144	72	90	4.1	1.84E+22
31	28/11/2019	04:23	41.465	19.633	22	154	62	94	4.6	8.80E+22
32	29/11/2019	08:02	41.547	19.640	25	142	73	89	4.0	1.48E+22
33	30/11/2019	20:53	41.568	19.524	25	150	74	86	4.5	6.07E+22
34	01/12/2019	06:04	41.586	19.590	25	160	77	84	4.2	2.72E+22
35	01/12/2019	11:42	41.342	19.464	28	135	65	68	3.9	9.80E+21
36	02/12/2019	02:08	41.488	19.697	24	152	70	84	4.4	5.12E+22

To benchmark stress orientations and stress-ratio proxies against the long-term regional field, we assembled a regional subset of 27 mechanisms (1964–2019) found in (Muço, 1994; Baker *et al*. 1997; Muço, 2007; Kiratzi and Dimakis, 2013; Roumelioti *et al*. 2011; Dushi *et al*., 2014; Dushi *et al*., 2018; Ormeni *et al*., 2022). This temporal window (1964–2019) encompasses the full span of instrumental and published focal-mechanism data available for the region, beginning with the earliest catalogue of Muço (1994) and extending through the recent broadband and moment-tensor compilations. For sensitivity, the same event list was inverted under two deterministic plane assignments: (i) the active/preferred movement plane when indicated by the source, and (ii) the auxiliary (conjugate) plane. Event selection, magnitudes, depths, and spatial sampling were kept identical for both assignments. Together, Tables 1–2 and Figure 1 provide the data foundation for the coseismic TENSOR solution, the early post-seismic StressInverse result, and the regional comparison.

3. METHODOLOGY

We used only the published focal mechanisms and moment tensors described in DATA; no waveform re-inversions were performed. All strike–dip–rake (SDR) inputs were harmonized using the FMC workflow (Alvarez-Gomez, 2019). For each nodal plane, we computed its pole, slip vector from the reported rake, and stored them as unit direction cosines. Agency “active/preferred plane” flags were preserved; otherwise, both planes were retained with corresponding weights. Before inversion, we

applied the Win-Tensor orientation statistics to the mainshock nodal planes to define coherent subsets for TENSOR analysis.

For each subset, we calculated the concentration index (*Rm;* resultant length per total weight), the Fisher-type concentration parameter *KA* = (n − 2)/ (n − Rm), and the mean cone angle *MCA* = arccos (Rm). To identify girdle distributions, we examined intersection B-axes following Lisle and Leyshon (2004)**.** According to Delvaux and Sperner (2003), a dataset was classified as "cluster-type" when pole statistics showed higher concentration than the corresponding B-axes. A minimum intersection angle of max (10°, MCA/2) was imposed to exclude nearly parallel pairs, and Win-Tensor weighting modes were applied (Mode 1: *w*, Mode 2: w^2, Mode 3: 10 *w*). In our dataset, this procedure isolated two stable FP1 clusters, with dominant strikes of approximately 150–155° and 337–352° (Fig. 3).

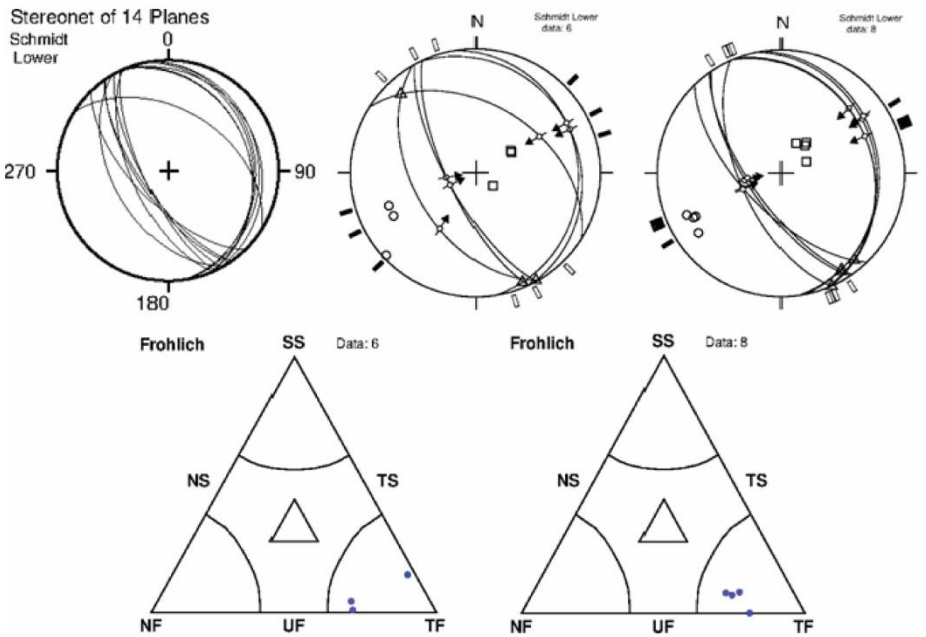


Fig.3. Reported active planes (FP1) for the 26 Nov 2019 Durrës mainshock (Mw 6.4): **a)** combined FP1 set (n = 14); **b)** subset-1 (n = 6); **c)** subset-2 (n = 8). Panels show lower-hemisphere Schmidt stereonets (top row) and Frohlich diagrams (bottom row) demonstrating a thrust (TF) regime. Sources: agency solutions summarized in **Table 1**.

The mainshock tensor was solved with TENSOR (Delvaux and Sperner, 2003), which combines Right-dihedron logic with rotational optimization to minimize slip-angle misfit (Fig. 4). Inputs consisted of the multi-agency SDRs listed in Table 1 after the subset selection above. TENSOR outputs principal stress orientations (σ1, σ2, σ3), the stress shape ratio φ = (σ2- σ3)/ (σ1- σ3), and quality metrics (QRw, aggregated as QRt). Separate inversions were performed for each FP1 subset, as well as a combined, cluster-aware inversion. The combined result is reported here, with subset results as sensitivity analysis.

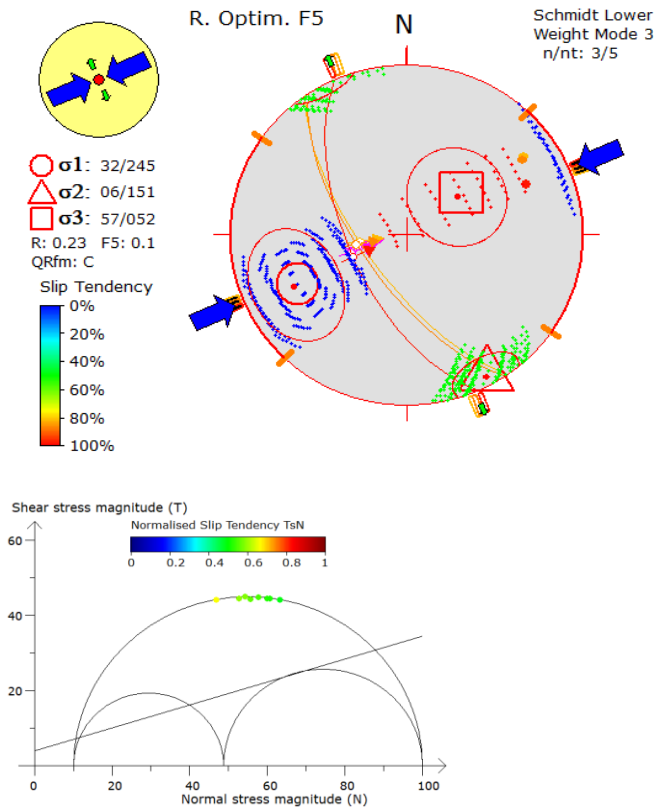


Fig. 4. Coseismic TENSOR solution for the 26 November 2019 Durrës mainshock. (a) Right-dihedron (Schmidt, lower hemisphere) projection showing FP1 planes from Table 1 with slip-tendency shading; principal stress axes (σ_1, σ_2, σ_3) and quality class are annotated. (b) Mohr diagram (shear vs. normal stress in MPa, plotted with consistent scale) confirming a thrust-fault (TF) type and a compressional stress regime.

The stress tensor for the seismic sequence (mainshock and 36 aftershocks; Table 2) was estimated using the Python implementation of *StressInverse* (*Vavryčuk, 2014*), based on the *Wallace–Bott assumption* (*Bott, 1959*), which states that slip occurs parallel to the direction of maximum resolved shear traction on a fault plane. The inversion minimizes the weighted angular misfit between observed slip directions and modeled shear stress to determine the principal stress axes (σ_1, σ_2, σ_3), the stress shape ratio φ (as defined above), and the apparent friction coefficient μ.

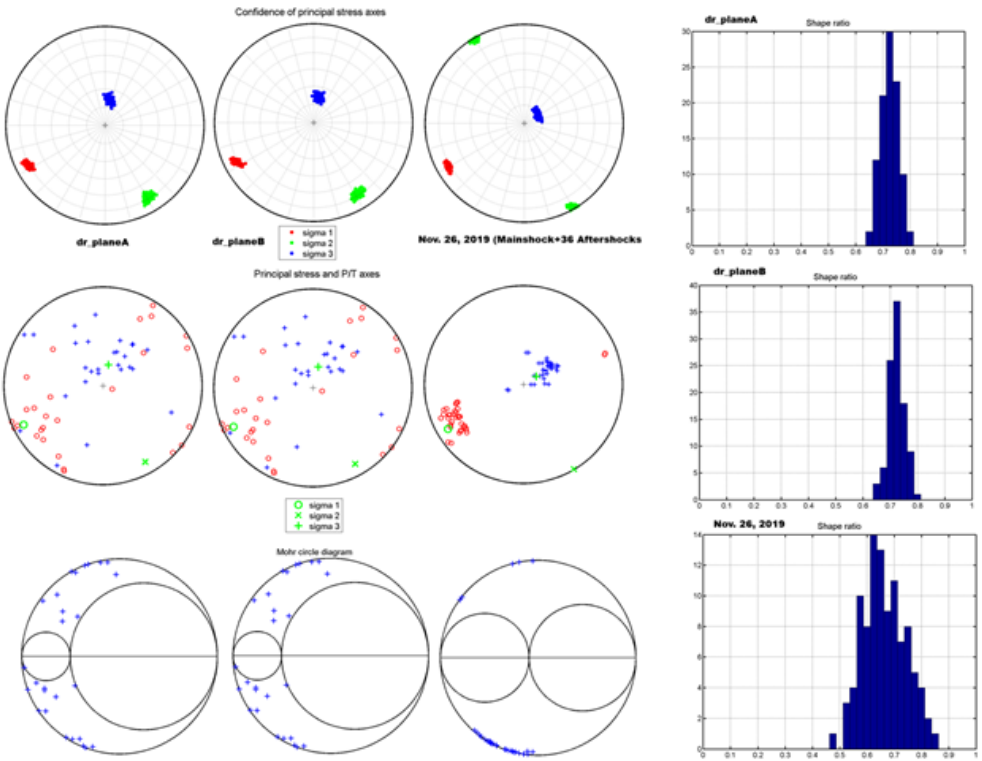


Fig. 5. StressInverse results for three cases: movement-plane policy, auxiliary-plane policy, and the 26 Nov 2019 sequence (mainshock + 36 aftershocks). Top: confidence regions of principal stress axes on lower-hemisphere stereonets (σ_1 red, σ_2 green, σ_3 blue). Middle: principal stress and P/T axes. Bottom: Mohr diagrams (normal vs. shear stress, in MPa) and histograms of the stress-shape ratio φ. Plane-policy definitions are applied consistently across datasets.

Agency-flagged active planes were used when available; otherwise, both planes were admitted with reduced weights, and a B-axis filter was applied to remove near-duplicate geometry. Dispersion and uncertainty was summarized using R_m, KA, and MCA on the resulting σ-axis populations. Cross-method consistency was ensured by comparing σ-axis bands, expressing TENSOR's ratio in the same φ convention as StressInverse, and verifying that dispersion and misfit metrics fell within expected ranges for reverse-faulting solutions.

4. RESULTS

Stress inversions were evaluated across three windows: (i) a regional benchmark of 27 published mechanisms (tested under two deterministic plane policies, active-plane and auxiliary-plane), (ii) a coseismic tensor for the mainshock (multi-agency focal-plane ensemble) using TENSOR, and (iii) an early post-seismic tensor using StressInverse (mainshock plus 36 immediate aftershocks). Numerical results are listed in Table 3; stereonets, P/T distributions, and Mohr plots are shown in the Figures 3 to 5.

Regional stress inversions yield compression-dominated tensors characterized by shallow σ_1 and steep σ_3 plunges. Under the active-plane policy, σ_1 = 246.9/8.1, σ_2 = 155.4/10.4, σ_3 = 14.0/76.7, with φ = 0.75 and μ = 0.95. Under the auxiliary-plane policy, σ_1 = 248/6.9, σ_2 = 151.0/11.0, σ_3 = 16.8/77.4, with φ = 0.70 and μ = 0.93. The σ_1 azimuth differs by ~1°, the confidence regions overlap (Fig. 5), indicating a stable WSW–ENE shortening direction weakly sensitive to plane assignment. The coseismic solution from TENSOR for the mainshock ensemble gives σ_1 = 244/14, σ_2 = 152/26, σ_3 = 26/70 (Fig. 4). This configuration is compatible with reverse-fault kinematics on ENE-dipping ramps beneath the coastal plain and consistent with the mechanisms reported by various contributing agencies.

The early post-seismic phase (mainshock and 36 aftershocks) yields σ_1 = 239.5/12.5, σ_2 = 149.3/0.49, σ_3 = 57.2/77.5, with φ = 0.546 and μ = 0.45 (Fig. 5). Although φ and μ decrease, indicating a more spheroidal stress ellipsoid, the P/T axes appear visibly more clustered. This reflects the kinematic coherence of the immediate aftershock sequence, where events rupture similar structures, even though the eigenvalue spacing of the stress tensor becomes temporarily reduced. The lower μ values, while not representing friction directly, are consistent with a transient reduction in differential stress, an effect commonly interpreted as apparent fault

weakening immediately after the mainshock (Dieterich, 1979; cf. Scholz, 2002; Delvaux and Sperner, 2003). Across all three windows, the stress regime remains consistently compressional (thrust-fault type), with σ_1 oriented WSW–ENE (≈239–248°, plunge ≈7–14°) and σ_3 steeply plunging (≈70–78°). (Table 3; Figs. 3–5).

This uniform orientation across regional, coseismic, and early post-seismic analyses demonstrates that the stress field preserved its compressional character throughout**,** despite short-term variations in φ and μ. The drop in φ and μ during the early post-seismic stage reflects a temporary reshaping of the stress field following the mainshock, consistent with a short-term reduction in differential stress. Although these parameters do not quantify frictional strength, their evolution aligns with the expected stress relaxation and apparent weakening of the fault system immediately after rupture. These transient adjustments occur within a long-lived compressional regime driven by Adria–Eurasia convergence, where periodic slip on adjacent thrust structures accommodates regional shortening (Fig. 1).

The coseismic TENSOR solution lies between the two regional variants and the early post-seismic solution, maintaining a reverse-faulting (compressional) regime. Concordance between StressInverse (regional and sequence windows) and TENSOR (coseismic window), combined with multi-agency mainshock inputs, confirms that the compressional stress field is consistently resolved by independent methods. These results align with prior structural interpretations (e.g., Teloni *et al*. 2021), which document WSW–ENE maximum compression and reverse faulting, and extend them by explicitly quantifying the early aftershock adjustment and the associated rise in φ and reduction in μ.

Table 3. Stress-tensor solutions by window and method. σ1–σ3 in azimuth/plunge (°); φ = stress-shape ratio; μ = apparent friction. N_events counts mechanisms (post-seismic = mainshock + 36). RO = rotational optimization; "—" = not reported

Window	**Method**	**No. events**	**σ1 (Az/Pl°)**	**σ2 (Az/Pl°)**	**σ3 (Az/Pl°)**	**φ**	**M**
Regional (active-plane policy)	StressInverse	27	246.9 / 8.1	155.4 / 10.4	14.0 / 76.7	0.75	0.95
Regional (auxiliary-	StressInverse	27	248.0 / 6.9	151.0 / 11.0	16.8 / 77.4	0.70	0.93

plane policy)							
Coseismic mainshock	TENSOR (RO)	1	244 / 14	152 / 26	26 / 70	—	—
Post-seismic sequence (26 Nov–2 Dec 2019)	StressInverse	37*	239.5 / 12.5	149.3 / 0.5	57.2 / 77.5	0.546	0.45

Notes: N_events = mechanisms used (post-seismic = mainshock + 36).

Table 4. Results obtained by the rotational optimization procedure (TENSOR).

Shear (incl/azm)	Sense	alpha (deg)	Tsn	M-dist	Nmag	Tmag	F5	Type	Comment
64/278	IS	7.7	0.55	13.6	56.9	44.5	2.2	R F	MT-INGV
63/274	IS	7.0	0.58	10.8	54.3	44.6	1.8	R F	MT-GCMT
62/261	IS	6.6	0.63	5.6	49.5	44.7	1.4	R F	MT-USGS

Table 5. Associated kinematic axes and map parameters.

(°)P incl (degrees)	azm (°)P azm (degrees)	(°)B incl (degrees)	azm (°)B azm (degrees)	(°)T incl (degrees)	azm (°)T azm (degrees)	(°)SHmax (degrees)	(°)SHmin (degrees)	R'	Reg.
24	243	9	148	64	39	69	159	2.5	TF
22	243	8	149	66	40	68	158	2.5	TF
18	246	0	156	72	65	66	156	2.5	TF

Table 6. Quality of the TENSOR solution.

Wms quality rank (QRw)	n	n/nt	CLW	alpha_w (deg)	DTW	P_len	S_len	Tensor quality rank (QRt)
B	8	0	0.5	3.8	0.75	0.71	0.73	B

Notes: incl/azm = inclination (plunge) / azimuth (deg); alpha = angular misfit; Tsn = normalized slip tendency; M-dist = misfit distance; Nmag/Tmag = normal and shear stress magnitudes (arbitrary units); F5 = stress ratio parameter; R' = inverse stress ratio; Reg. = faulting regime; CLW, alpha_w, DTW, P_len, S_len follow Win-Tensor/WSM conventions.

Based on the reduced stress tensor, the principal stresses are σ_1 = 23/249 (±24.3), σ_2 = 05/157 (±12.8), and σ_3 = 66/056 (±24.6), with R = 0.5 and R′ = 2.5, and a maximum horizontal stress oriented ~ 69/159. This configuration corresponds to a thrust-fault (TF) regime and a compressional stress field, in full agreement with the moment-tensor solutions from GCMT, INGV, and USGS. The slip-tendency and misfit parameters derived from the three independent moment-tensor sources (MT-INGV, MT-GCMT, MT-USGS) also converge on a reverse-faulting pattern, with low angular misfits ($\alpha \approx 6.6$–$7.7°$), moderate normalized slip tendency (Tsn ≈ 0.55–0.63), and stress magnitudes consistent across all datasets. The P/B/T axes extracted from Win-Tensor further confirm a stable thrust-fault orientation (R′ = 2.5), reinforcing the interpretation of a regionally coherent compressional regime affecting the Outer Albanides.

5. DISCUSSION

The stress inversions for the regional, coseismic, and early post-seismic windows consistently define a compressional, thrust-fault (TF) stress regime, in agreement with Teloni et al. (2021). Using Win-Tensor and rotational optimization, those authors obtained for the 2019–2020 Durrës sequence a stress ratio R = 0.5 and principal axes σ_1 = 15°/241°, σ_2 = 10°/148°, and σ_3 = 72°/025°, indicating a WSW–ENE maximum compression and a near-vertical minimum stress. Their solution, ranked B in the World Stress Map classification (Heidbach et al., 2016), reflects a stable reverse-faulting regime associated with regional Adria–Eurasia shortening.

The present results reproduce this orientation closely ($\sigma_1 \approx 239$–$248°$, plunge ≈ 7–14°; $\sigma_3 \approx 70$–$78°$) and extend the interpretation by explicitly quantifying the stress-shape ratio φ and the stress-ratio parameter μ, which were not evaluated in earlier studies of this sequence. Following the classification schemes of Delvaux & Sperner (2003) and Vavryčuk (2014), the regional tensor ($\varphi \approx 0.70$–0.75, $\mu \approx 0.93$–0.95) corresponds to a well-defined triaxial compressional state typical of long-term thrust-fault provinces.

The early post-seismic solution ($\varphi \approx 0.546$, $\mu \approx 0.45$) exhibits a more oblate stress ellipsoid and a marked reduction in differential stress, while the orientation of σ_1 and σ_3 remains unchanged. This shift from a triaxial to a more oblate configuration captures the short-term mechanical response of the crust immediately after the mainshock. As illustrated by the step-wise coseismic stress-relaxation pattern (Fig. 6), this transition reflects a

temporary decrease in stress anisotropy rather than a rotation of the stress field. Although μ does not represent true friction, its reduction, together with the diminished eigenvalue spacing, aligns with an apparent weakening of the activated thrust system during the early aftershock phase (Dieterich, 1979; Scholz, 2002).

Overall, the combined regional, coseismic, and post-seismic results indicate a robust, time-persistent compressional regime driven by Adria–Eurasia convergence, within which transient coseismic perturbations modulate the stress magnitudes and shape parameters without altering the long-term principal-stress orientation.

The temporal evolution of the Durrës sequence (Fig. 6) reveals a distinctly segmented pattern of energy release. Following low-level background activity in early 2019, a prolonged quiescent period (June–August) preceded the September foreshocks and the subsequent November–December mainshock–aftershock cluster. The Gaussian-smoothed magnitude curve highlights these discrete accelerations of activity, separated by intervals of near silence, confirming a discontinuous, stepwise sequence evolution. The cumulative-energy trend shows a sharp increase during the mainshock, followed by smaller stepwise rises due to ML 5.0–5.5 aftershocks, consistent with episodic stress release governed by transient reductions in μ (≈ 0.45). This sequential behavior aligns with the inversion results and supports a friction-controlled instability operating under a persistent compressional regime, as inferred from the regional stress field (Delvaux and Sperner 2003; Vavryčuk 2014; Teloni *et al.* 2021). The alternation between quiescence and reactivation indicates that stress redistribution occurred locally without significant reorientation of σ_1, implying deformation dominated by the regional Adria–Eurasia compression.

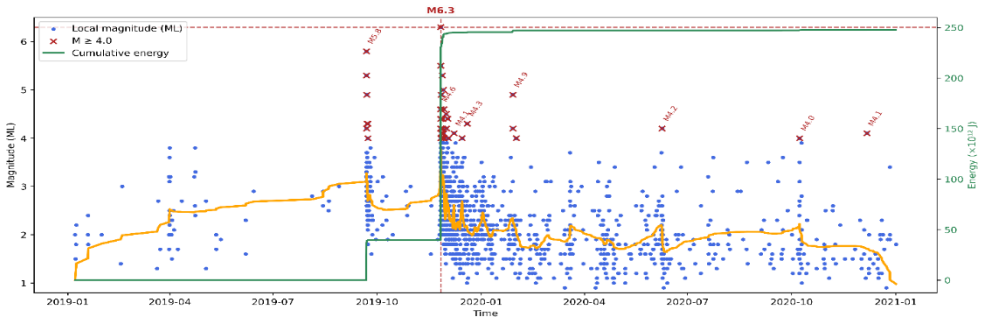


Fig. 6. Sequentially smoothed magnitude–time series and cumulative seismic energy release for the 2019–2021 Durrës sequence.

Across all inversion windows, the stress field remains compressional (thrust-fault type), with σ_1 oriented WSW–ENE (≈ 239–248°, plunge ≈ 7–14°) and σ_3 near vertical (≈ 70–78°). This persistence across the regional, coseismic, and post-seismic solutions shows that the stress field maintains a stable compressional character, while short-term variations in φ and μ reflect transient stress relaxation and apparent weakening within the same long-term regime. Such behavior accords with ongoing Adria–Eurasia shortening, where local stress adjustments are periodically accommodated by slip on nearby thrusts in the Outer Albanides (Fig. 1). The agreement between StressInverse (regional and sequence analyses) and TENSOR (coseismic analysis) confirms the robustness of the compressional field resolved by independent datasets and methods, consistent with regional stress compilations from western Greece and the southern Adriatic (e.g., Kiratzi and Dimakis, 2013; Roumelioti *et al*. 2011).

The StressInverse computations used the program's default azimuth–plunge grid (5° step), which proved sufficient to reproduce consistent thrust-faulting tensors across all analysis windows. No further refinement was required. The TENSOR solution was similarly refined through rotational optimization, improving focal-plane orientation without altering the faulting style. The convergence of both methods toward identical σ_1 orientation and stress regime confirms that the adopted configurations are robust and consistent with the recommended implementations of Delvaux & Sperner (2003) and Vavryčuk (2014).

6. CONCLUSIONS

The Durrës–Tirana sector is characterized by a persistent WSW–ENE compressional regime consistent with Adria–Eurasia convergence and the mapped thrust architecture of the Outer Albanides. Across all inversion windows, σ_1 trends ~239–248° with a shallow plunge (~7–14°), and σ_3 remains near vertical (~70–78°), defining a thrust-fault (TF) stress regime.

Regional inversions of 27 published mechanisms (φ ≈ 0.70–0.75, μ ≈ 0.93–0.95) reveal a well-defined triaxial compressional state. In contrast, the coseismic tensor for the 26 November 2019 mainshock (σ_1 = 244/14, σ_2 = 152/26, σ_3 = 26/70) and the early post-seismic inversion (φ ≈ 0.546, μ ≈ 0.45) document transient stress relaxation and apparent weakening immediately after rupture. The φ–μ evolution from a higher-stress regional state to a temporarily more oblate post-seismic configuration parallels the observed step-wise clustering of seismicity and cumulative-energy release,

indicating short-term stress redistribution followed by gradual re-strengthening.

Independent analyses using both TENSOR and StressInverse converge on a robust, thrust-dominated stress field, capturing these subtle but short-lived adjustments. The results suggest that deformation within the Durrës–Tirana corridor is primarily governed by ongoing Adria–Eurasia shortening, with transient stress transfer and partial reloading along nearby thrusts modulating post-seismic evolution. This dual-method approach strengthens confidence in the resolved stress tensors and provides a solid framework for future assessments of stress transfer and strain partitioning across the active Albanides.

ACKNOWLEDGMENTS

A preliminary version of this analysis was presented as a poster at the *12th Congress of the Balkan Physical Union (BPU-12)*, held in Bucharest, Romania, from 8 to 12 July 2025, and hosted by the National University of Science and Technology POLITEHNICA Bucharest (Poster Session 2/159; *Book of Abstracts*, p. 101).

The author gratefully acknowledges the constructive comments and critical insights of the anonymous reviewer, whose valuable suggestions substantially improved the clarity and consistency of this manuscript.

REFERENCES

Aliaj Sh. 2016. *Tectonics and Geodynamics of Albania.* Tirana: Academy of Sciences of Albania.

Álvarez-Gómez JA. 2019. FMC—Earthquake focal mechanisms data management, cluster and classification. *SoftwareX*, 9, 299–307. https://doi.org/10.1016/j.softx.2019.03.008.

Baker C, Hatzfeld D, Lyon-Caen H, Papadimitriou E, Rigo A, Deschamps A. 1997. Earthquake mechanisms of the Adriatic Sea and western Greece. *Geophysical Journal International*, **131**, 559–594. https://doi.org/10.1111/j.1365-246X.1997.tb06600.x.

Bott MHP. 1959. The mechanics of oblique-slip faulting. *Geological Magazine*, **96 (2):** 109–117. https://doi.org/10.1017/S0016756800059987.

D'Agostino N, Avallone A, Cheloni D, D'Anastasio E, Mantenuto S, Selvaggi G. 2008. Active tectonics of the Adriatic region from GPS

and earthquake slip vectors. *Journal of Geophysical Research: Solid Earth*, **113**, B12413. https://doi.org/10.1029/2008JB005860.

Delvaux D, Sperner B. 2003. New aspects of tectonic stress inversion with reference to the TENSOR program. In D. A. Nieuwland (Ed.), *New Insights into Structural Interpretation and Modelling* (pp. 75–100). Geological Society, London, Special Publications 212. https://doi.org/10.1144/GSL.SP.2003.212.01.06.

Dieterich JH. 1979. Modeling of rock friction: 1. Experimental results and constitutive equations. *Journal of Geophysical Research,* 84, 2161–2168. https://doi.org/10.1029/JB084iB05p02161.

Drakatos G, Voulgaris N, Pirli M, Melis N, Karakostas B 2005. 3-D crustal velocity structure in northwestern Greece. *Pure and Applied Geophysics*, **162**, 37–51. https://doi.org/10.1007/s00024-004-2578-2.

Dushi E, Koçi R, Begu E, Bozo Rr. 2018. Stress inversion from focal mechanisms of moderate earthquakes in Albania. *International Multidisciplinary Scientific GeoConference SGEM*, **18(1.1)**: 989–996. https://doi.org/10.5593/sgem2018/1.1/S05.123.

Dushi E, Minarolli A, Kasaj E, Gjuzi O. 2014. Focal mechanism solutions for local earthquakes (M > 3.0) from Albanian Seismological Network broadband recordings. In *XX Congress of the Carpathian-Balkan Geological Association (CBGA)*, Tirana, Vol. I, Session G05.

EIDA (ORFEUS). European Integrated Data Archive. Available at http://www.orfeus-eu.org/data/eida/

EMSC (Euro-Mediterranean Seismological Centre). (2024). *Moment tensor database and earthquake information service.* Available at https://www.emsc-csem.org/Earthquake_data/tensors.php.

Ganas A, Elias P, Briole P, Cannavo F, Valkaniotis S, Tsironi V, Partheniou EI. 2020. Ground deformation and seismic fault model of the M6.4 Durrës (Albania) 26 Nov 2019 earthquake based on GNSS/INSAR observations. *Geosciences*, **10**, 210. https://doi.org/10.3390/geosciences10060210.

Heidbach O, Rajabi M, Reiter K, Ziegler M. 2016. *World Stress Map database release 2016.* GFZ Data Services. https://doi.org/10.5880/WSM.2016.001.

Jouanne F, Mugnier, J-L, Koci R, Bushati S, Matev K, Kuka N, Pérouse E. 2012. GPS constraints on current tectonics of Albania. *Tectonophysics*, **554–557**, 50–62. https://doi.org/10.1016/j.tecto.2012.06.008.

Kiratzi A, Dimakis E. 2013. Focal mechanisms and slip models of moderate-size earthquakes in Albania and adjacent countries. *Italian Journal of Geosciences*, **132(2):** 186–193. https://doi.org/10.3301/IJG.2011.33.

Kiratzi AA, Louvari E. 2003. Focal mechanisms of shallow earthquakes in the Aegean Sea and the surrounding lands determined by waveform modelling: A new database. *Journal of Geodynamics*, **36**, 251–274. https://doi.org/10.1016/S0264-3707(03)00050-4.

Konstantinou KI, Melis NS, Boukouras K. 2010. Routine regional moment tensor inversion for earthquakes in the Greek region: The NOA database (2001–2006). *Seismological Research Letters*, **81(5):** 750–760. https://doi.org/10.1785/gssrl.81.5.750.

Kristina Matraku, François Jouanne, Edmond Dushi, Rexhep Koçi, Neki Kuka, Raphael Grandin, Pascale Bascou. 2023. The 26 November 2019 Durrës earthquake, Albania: Coseismic displacements and occurrence of slow-slip events in the year following the earthquake. *Geophysical Journal International*, **234**(2): 807–838. https://doi.org/10.1093/gji/ggad101.

Lekkas E, Mavroulis S, Papa D, Carydis P. 2019. The 26 November 2019 Mw = 6.4 Durrës (Albania) earthquake. *Newsletter on Environmental Disaster and Crisis Management Strategies*, **15**, 80. ISSN 2653-9454.

Lisle RJ, Leyshon PR. 2004. *Stereographic Projection Techniques for Geologists and Civil Engineers* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139171366.

Matraku K, Dushi E, Muceku Y, Ganas A, Rama, B. 2023. The 26 November 2019 Durrës earthquake, Albania: Coseismic displacements and occurrence of slow-slip events in the year following the earthquake. *Geophysical Journal International*, **234(2):** 807–838. https://doi.org/10.1093/gji/ggad101.

Mavroulis S, Lekkas E, Carydis P. 2021. Liquefaction phenomena induced by the 26 November 2019, Mw = 6.4 Durrës (Albania) earthquake and liquefaction susceptibility assessment in the affected area. *Geosciences*, **11(5):** 215. https://doi.org/10.3390/geosciences11050215.

Métois M, Vigny C, Chéry J, Huchon P, Ruegg J-C, Briole P. 2015. Insights on continental collisional processes from GPS data: Dynamics of the peri-Adriatic belts. *Journal of Geophysical*

Research: Solid Earth, **120(11):** 8701–8720. https://doi.org/10.1002/2015JB012023.

Michael AJ. (1984): Determination of stress from slip data: Faults and focal mechanisms. *Journal of Geophysical Research: Solid Earth*, **89(B13),** 11517–11526. https://doi.org/10.1029/JB089iB13p11517.

Michael AJ. 1987. Use of focal mechanisms to determine stress: A control study. *Journal of Geophysical Research: Solid Earth*, **92(B1),** 357–368. https://doi.org/10.1029/JB092iB01p00357.

Moshou A, Dushi E. 2019. A preliminary report on the 26 November 2019 Mw = 6.4 Durrës, Albania earthquake. European-Mediterranean Seismological Centre, Essonne, France, 12 pp.

Muço B. 2007. Focal mechanism solutions and stress field distribution in Albania. *Albanian Journal of Natural & Technical Sciences*, **1**, 129–134.

Ormeni, R, Hoxha I, Gjuzi O, Bozo R, Gega D, Kanani Xh, Mucaj D, Piccardi L, Vittori E, Blumetti AM, Di Manna P, Comerci V. 2022. The catalogue of earthquake focal mechanisms in Albania and surroundings (1948–2022). *EAGE/EarthDoc Extended Abstracts.* https://doi.org/10.3997/2214-4609.202220159.

Papadopoulos GA, Agalos A, Carydis P, Lekkas E, Mavroulis S, Triantafyllou I. 2020. The 26 November 2019 Mw = 6.4 Albania destructive earthquake. *Seismological Research Letters*, **91 (6):** 3129–3138. DOI:10.1785/0220200207.

Randall GE. 1994. Efficient computation of Green's functions in layered media using the reflectivity method. *U.S. Geological Survey Open-File Report 94-XX.*

Roumelioti Z, Kiratzi A, Benetatos C. 2011. Time-domain moment tensors for shallow earthquakes in the broader Aegean Sea (2006–2007). *Journal of Geodynamics*, **51 (2-3)**, 179–189. https://doi.org/10.1016/j.jog.2010.01.011.

Sokos EN, Zahradník J. 2008. ISOLA: A Fortran code and a MATLAB GUI to perform multiple-point-source inversion of seismic data. *Computers & Geosciences*, **34(8):** 967–977. https://doi.org/10.1016/j.cageo.2007.07.005.

Teloni S, Invernizzi C, Mazzoli S, Pierantoni PP, Spina V. 2021. Seismogenic fault system of the Mw 6.4 November 2019 Albania earthquake: New insights into the structural architecture and active tectonic setting of the Outer Albanides. *Journal of the Geological Society*, **178(2):** jgs2020-193. https://doi.org/10.1144/jgs2020-193.

Triantafyllis N, Ilias A, Sokos EN. 2022. GISOLA: A high-performance computing application for real-time moment tensor inversion. *Seismological Research Letters*, **93(2A):** 957–968. https://doi.org/10.1785/0220210153.

Vavryčuk V. 2014. Iterative joint inversion for stress and fault orientations from focal mechanisms. *Geophysical Journal International*, **199(1)**: 69–77. https://doi.org/10.1093/gji/ggu224.

Wiemer S. 2001. A software package to analyze seismicity: ZMAP. *Seismological Research Letters*, **72(3):** 373–382. https://doi.org/10.1785/gssrl.72.3.373.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

FROM MONO-CLASS TO MULTI-CLASS PHYSICS-GUIDED TRAFFIC STATE PREDICTION: A COMPARATIVE STUDY OF GAUSSIAN PROCESSES AND LSTM NETWORKS

Kleona BINJAKU, Elinda KAJO MEÇE
Department of Computer Engineering, Polytechnic University of Tirana, Albania

Cecilia PASQUALE and Simona SACONE
Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Italy

Author for correspondence: Kleona.Binjaku@fti.edu.al

KB, 0009-0006-9019-2481; EKM, 0009-0002-7353-476;
CP, 0000-0002-0760-4769; SS, 0000-0002-7114-8972
https://doi.org/10.67537/ajnts3001004

ABSTRACT

Accurate traffic prediction is fundamental to effective urban mobility. This paper investigates the shift from mono-class to multi-class traffic modelling and its implications for physics-guided machine learning techniques. We examine two approaches for multi-class highway traffic prediction, encompassing both passenger vehicles and trucks: (i) a physics-regularized Gaussian Process (GP) model that integrates constraints from the multi-class METANET model, and (ii) a Long Short-Term Memory (LSTM) network that incorporates these constraints into its loss function. Beyond evaluating prediction accuracy, computational cost, and implementation complexity, we highlight the challenges introduced by multi-class formulations—particularly the need to account for diverse vehicle dynamics. Our findings provide insights into the trade-offs and feasibility of each approach and assess the impact of transitioning from mono-class to multi-class models on predictive performance.

***Keywords*:** Gaussian Processes, machine learning, LSTM networks, multi-class METANET model, traffic modelling

1. INTRODUCTION

Accurate evaluation and prediction of traffic conditions are essential for improving traffic flow management, reducing congestion, and enhancing safety in Intelligent Transportation Systems (ITS). Traditional

physics-based models, such as the METANET model (Papageorgiou *et al*., 1989), offer interpretable representations of macroscopic traffic dynamics. However, they often require extensive calibration and struggle to capture the stochastic nature of real-world traffic. Conversely, machine learning models have demonstrated strong predictive capabilities by leveraging large datasets, yet they may lack physical consistency and generalizability—especially in scenarios with limited data. To address these challenges, we propose two physics-guided machine learning models that integrate the strengths of data-driven approaches with domain knowledge derived from the widely adopted multi-class METANET framework.

The multi-class METANET model (Pasquale *et al*. 2015), an extension of the original METANET framework, incorporates distinct dynamics and interaction terms for different vehicle classes, enabling a more accurate representation of mixed traffic flow. This enhances the realism and precision of traffic state modeling, particularly in scenarios involving heterogeneous traffic composed of both passenger vehicles and trucks.

In this paper, we build upon our previous work on mono-class traffic prediction (Binjaku *et al*. 2024) by extending it to a multi-class context that explicitly accounts for vehicle heterogeneity. In our earlier study, we compared physics-regularized Gaussian Processes and physics-informed LSTM networks within a mono-class traffic setting. However, since real-world traffic comprises multiple vehicle types exhibiting diverse behaviours, mono-class models may fall short in accurately capturing these complex dynamics.

This study addresses two key research questions:

- What are the benefits of extending physics-guided machine learning models from a mono-class to a multi-class traffic setting that explicitly captures vehicle heterogeneity?
- How does the performance of a physics-informed Long Short-Term Memory (LSTM) model compare to that of a physics-regularized Gaussian Process (GP) when both are applied to multi-class traffic prediction?

To answer these questions, we propose and evaluate two multi-class physics-guided models based on the METANET framework:

1. A multi-class physics-regularized Gaussian Process, in which the equations from the multi-class METANET model are incorporated as

a regularization term to guide the GP in learning traffic dynamics, while guaranteeing adherence to basic traffic flow relationships.

2. A multi-class physics-informed Long Short-Term Memory network, which enhances the ability of an LSTM neural network to model temporal dependencies by directly embedding traffic flow physics into its architecture, thereby preserving physically meaningful constraints.

Both models explicitly consider two vehicle classes—cars and trucks—allowing for a more accurate presentation of mixed traffic flow, both of which are critical indicators for evaluating and managing traffic conditions. are explicitly taken into account in both models, enabling a more accurate depiction of mixed traffic flow dynamics. The primary traffic variables predicted are mean speed and traffic flow, both of which are critical indicators for evaluating and managing traffic conditions.

To rigorously evaluate the performance of the proposed models, we compare them across three key dimensions: i) accuracy of predictions, ii) training time, and iii) complexity of implementation.

We perform and a comparative analysis of the physics-informed LSTM and the physics-regularized Gaussian Process models in both mono- and multi-class settings. The inclusion of results from our prior mono-class framework enables a comprehensive and consistent evaluation, allowing us to assess the advantages of incorporating multiple vehicle classes in modelling process.

The reminder of the paper is structured as follows: Section 2 provides an overview of related work on the METANET model, Gaussian Processes, LSTMs, and physics-guided machine learning models for traffic analysis. Section 3 introduces the proposed multi-class physics-based machine learning models and summarizes all models employed in this study. Section 4 details the experimental setup using real-world traffic data. Finally, Section 5 presents the conclusions and outlines directions for future research.5.

2. RELATED WORK

The literature on traffic modelling broadly categorizes existing approaches into three groups: data-driven models, physics-based models, and hybrid approaches that integrate domain knowledge into machine learning frameworks. Traditional physics-based models—such as microscopic and macroscopic traffic models—have been widely adopted

over the years. Macroscopic models offer a higher-level representation of traffic flow dynamics and include first-order models like the Lighthill-Whitham-Richards (LWR) model, as well as second-order models such as Payne-Whitham (PW) and Aw-Rascle-Zhang (ARZ). Among these, the METANET model has gained particular prominence due to its discrete-time, discrete-space formulation, which enables efficient traffic state estimation and supports practical control applications. Numerous extensions of the METANET framework have been proposed to handle heterogeneous vehicle classes (Pasquale *et al*. 2014; 2015), autonomous and connected vehicles (Shahri *et al*. 2022), and freeway control systems incorporating features such as ramp metering and variable speed limits (Hegyi *et al*. 2005; Chavoshi *et al*. 2023). These enhancements have improved the model's flexibility and applicability in addressing the demands of modern traffic management.

With the increasing availability of traffic data, data-driven models have become powerful tools for estimating and forecasting traffic conditions. Among these, Gaussian Process (GP) models have been widely adopted due to their probabilistic nature and inherent ability to quantify uncertainty. Applications of GP models span a range of tasks, including scalable predictions via infinite mixture models (Sun and Xu, 2011) and queue length forecasting (Kocijan and Přikryl, 2010). Their predictive performance has been further improved through hybrid approaches such as GP ensembles (Zhan *et al*. 2018) and GP-LSTM combinations (Xie *et al*. 2021). On the other hand, Long Short-Term Memory (LSTM) networks are particularly well-suited for capturing temporal dependencies in traffic data. They have demonstrated strong performance in various applications, including multi-step traffic forecasting (Liyong and Vateekul, 2019), imputation of missing data (Tian *et al*. 2018), and graph-based modelling for road link optimization (Liu *et al*. 2020). Moreover, the integration of Convolutional Neural Networks (CNNs) with LSTMs has enhanced the ability to detect spatiotemporal patterns in traffic networks (Chu *et al*. 2021; Lian and Wang 2024).

While physics-based models often struggle to adapt in real time, traditional data-driven models tend to overfit when faced with sparse or noisy data. Physics-guided machine learning (PGML) addresses these limitations by integrating physical laws into data-driven models, thereby enhancing the accuracy and robustness of real-time traffic management systems. PGML enables models to learn from data while remaining consistent with the underlying physical principles or domain knowledge of

the system being modelled. In recent years, interest in physics-guided machine learning has grown significantly. This increasing trend is illustrated in Figure 1, which is based on the body of literature identified during our review.

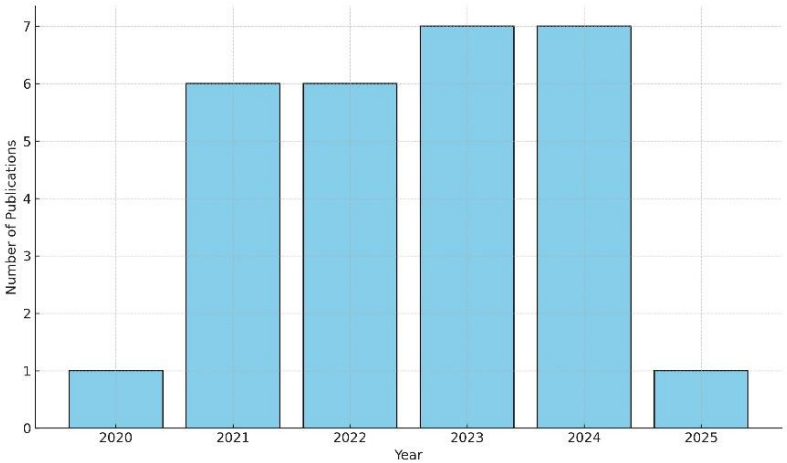


Fig.1: Distribution of publications about physics-based machine learning models for traffic state estimation.

Physics-guided machine learning (PGML) techniques have been applied to a wide range of tasks, from fuel consumption modelling and car-following behaviour to trajectory prediction. For instance, to reduce vehicle emissions, (Das and Tanvir 2024) integrates Vehicle-Specific Power into a physics-informed LSTM model for autonomous vehicles (AVs). To enhance trajectory prediction, several studies—such as (Long *et al*. 2024; Sheng *et al*. 2024), and (Geng *et al*. 2023) —incorporate traffic models like the Intelligent Driver Model (IDM) and vehicle kinematics into deep learning frameworks. In efforts to simulate driver behaviour, hybrid approaches also merge classical physics with machine learning techniques. Liu *et al.* (2023) employ quantile regression within a physics-informed deep learning (PIDL) framework to capture stochastic driving dynamics, while Mo *et al.* (2021) propose variations of PIDL-CF based on four distinct car-following models. Lei *et al*. (2024) utilize NSGA-II multi-objective optimization in conjunction with physical models to address both fuel consumption and battery degradation in plug-in hybrid electric vehicles (PHEVs).

Traffic State Estimation (TSE) has been the focus of several studies (Huang and Agarwal, 2020; 2022; 2023a,b; Huang *et al*. 2024). To address challenges such as sparse sensor coverage, noisy data, and complex boundary conditions, these studies iteratively refine their physics-informed deep learning (PIDL) models by incorporating traffic flow models such as the Lighthill-Whitham-Richards (LWR) model, the Cell Transmission Model (CTM), conservation laws, and nonlocal variants of the LWR model. In their later work, fog computing is introduced to improve the real-world applicability and scalability of the models.

Meanwhile, physics-informed neural networks have also been researched by other academics. Aw-Rascle-Zhang (ARZ) is a second-order traffic flow model that was incorporated into the PIDL architecture (Shi, 2021). Shi *et al*. (2022) enhance the model estimation performance under sparse data conditions by extending the PIDL framework with a Fundamental Diagram Learner (FDL). To more flexibly capture traffic dynamics, Di *et al*. (2023) propose a hybrid graph architecture that integrates both physics-informed and physics-uninformed components.

Other notable approaches include the LWR-based PIDL framework proposed by Rempe *et al.* (2021), which integrates multiple sources of traffic data, and the domain-decomposition-based Physics-Informed Neural Network (PINN) developed by Usama *et al*. (2022). Barreau *et al.* (2021a,b) utilize coupled micro–macro PINN models to jointly identify, reconstruct, and predict traffic states using sparse data collected from probe vehicles. Recent advancements include multi-task optimization techniques Wang *et al.* (2023), where auxiliary tasks are used to guide PINNs via the LWR model, and the work of Zhao and Yu (2022), who apply second-order partial differential equation (PDE) models to estimate traffic states both spatially and temporally.

Recurrent neural networks, particularly Long Short-Term Memory (LSTM) models, are frequently used in traffic state prediction due to their ability to capture temporal dependencies in traffic flow. The integration of physics-based traffic flow concepts has been shown to improve the accuracy and interpretability of LSTM models in several studies. For instance, Das and Tanvir (2024) present a physics-informed LSTM framework tailored for self-driving vehicles in mixed traffic conditions. The model employs a multi-objective loss function that incorporates Vehicle-Specific Power (VSP) to balance trajectory prediction with emission reduction. To enhance realistic car-following behaviour, Tischmann *et al.* (2024) combine LSTM with the Intelligent Driver Model

(IDM), using a physics-guided loss function to steer the training process. Fafoutellis and Vlahogianni (2025) employ Granger causality for feature selection and introduce a Traffic Flow Theory-Informed (TFTI) loss function to align LSTM predictions with fundamental traffic flow diagrams. In another study, Pereira *et al*. (2022) integrate the Traffic Reaction Model (TRM) into a physics-informed LSTM to ensure macroscopic consistency in traffic flow predictions.

Physics-guided reinforcement learning (RL) is also emerging as a promising approach in traffic management. Han *et al*. (2022) introduce a physics-informed RL ramp metering method that enhances prediction accuracy by integrating synthetic and real-world data, grounded in polynomial traffic flow models.

To achieve both flexibility and physical consistency, Sheng *et al*. (2024) propose a residual RL framework that embeds domain expertise via the Intelligent Driver Model (IDM) to model core traffic dynamics, while employing neural networks to learn residual patterns beyond those captured by traditional physics.

3. OUR PROPOSED PHYSICS-BASED MACHINE LEARNING MODELS

This section introduces the physics-based machine learning models proposed and developed in this study: the multi-class physics-regularized Gaussian Process (PR-GP) and the multi-class physics-informed Long Short-Term Memory (PI-LSTM) network. These models extend established approaches by integrating components from the Gaussian Process (GP) framework, the multi-class METANET traffic flow model, and Long Short-Term Memory (LSTM) networks, thereby combining data-driven learning with domain-specific traffic flow dynamics.

3.1 The physics-regularized Gaussian process

The multi-class physics-regularized Gaussian Process (PR-GP) model developed in this study is based on the integration of the multi-class METANET traffic flow model and Gaussian Processes (GPs). GPs are a powerful and flexible non-parametric framework used for regression and classification tasks, where relationships between data points are modelled probabilistically (Su and Zhang 2017). To ensure that the GP model remains consistent with fundamental traffic flow principles, a regularization term is incorporated into the objective function. This

regularization component is derived from the governing equations of the multi-class METANET model, initially introduced in (Pasquale *et al*., 2015), which extends the original METANET framework (Papageorgiou *et al*., 1989) by incorporating the dynamics of multiple vehicle classes. Specifically, the regularization term is formulated based on the multi-class METANET equations for traffic density, mean speed, and traffic flow (Equations 1, 2, and 3 in Pasquale *et al*., (2015)), and is computed as follows:

$$g_{1,c} = \widehat{\rho_{i,c}}(k+1) - \widehat{\rho_{i,c}}(k) - \frac{T}{L_i\lambda_i}\left[\widehat{q_{i-1,c}}(k) - \widehat{q_{i,c}}(k) + \widehat{r_{i,c}}(k) - \widehat{s_{i,c}}(k)\right] \quad (1)$$

$$g_{2,c} = v_{i,c}(k+1) - v_{i,c}(k) - \frac{T}{\tau_c}\left[V_{i,c}(k) - v_{i,c}(k)\right] - \frac{T}{L_i}v_{i,c}(k)\left(v_{i-1,c}(k) - v_{i,c}(k)\right) + \frac{\nu_c T(\rho_{i-1}(k) - \rho_i(k))}{\tau_c L_i\,(\rho_i(k) + \chi_c)} - \delta_{on}T\,\frac{v_{i,c}(k)r_i(k)}{L_i\,[\rho_i(k) + \chi_c]} \quad (2)$$

$$g_{3,c} = \widehat{q_{i,c}}(k) - \widehat{\rho_{i,c}}(k) \cdot \lambda_i \cdot \widehat{v_{i,c}}(k) \quad (3)$$

where $\hat{\rho}$, $\hat{v}, \hat{q}$ denote the estimated values of traffic density, speed, and flow, respectively.

Gaussian Processes are employed to estimate these quantities. To enhance the physical consistency of predictions, a set of pseudo-observations (Z,ω) is introduced. These pseudo-observations (Z,ω) have the same structure as the original input-output pairs (X,Y), where X represents the input features and Y the observed outputs. The pseudo-inputs Z are selected from the spatial-temporal domain to represent regions within data distribution. The corresponding pseudo- outputs ω are not directly, but are instead inferred from GP model and are used to evaluate adherence to physical constrains. The selection of pseudo-inputs Z is performed using a clustering-based approach, specifically the k-Means algorithm (Ikotun *et al*. 2023) to ensure good coverage of the spatial-temporal data domain.

Importantly, the pseudo-observations are selected to represent the input space in a realistic and data-representative manner. The locations Z are determined using a clustering-based technique, specifically the k-Means algorithm (Ikotun *et al*. 2023). This approach ensures that the pseudo-observations effectively cover the spatiotemporal domain, reflecting the underlying structure of the dataset. As a result, the model is able to evaluate and enforce physical constraints across a diverse range of traffic conditions.

At the selected pseudo-observation points, the model evaluates the predicted values of traffic density, speed, and flow, which are then used to assess the degree to which the multi-class METANET equations are satisfied. This evaluation enables the computation of physics-based residuals corresponding to each governing equation.

These residuals are used to derive the terms $g_{1,c}$, $g_{2,c}$, and $g_{3,c}$, which represent the deviations from the expected physical relationships for each vehicle class ccc. These terms are incorporated into the regularization strategy of the model. By embedding them into the Gaussian Process objective function, the model is guided to generate predictions that not only fit the observed data but also conform to the underlying physical traffic flow dynamics. This physics-based regularization, made possible through the use of pseudo-observations, enhances the model's generalization capabilities—particularly in regions with sparse or noisy data—while ensuring that the learned relationships remain consistent with established traffic flow theory.

Consistency with traffic flow dynamics is enforced through the inclusion of a physics-based regularization term, which enhances the Gaussian Process (GP) marginal likelihood. This augmented likelihood forms the basis of the final objective function. As a result, the model is trained not only to fit observed data but also to remain consistent with known physical laws governing traffic dynamics.

The posterior distribution of the GP model, incorporating both observational data and physics-based constraints, can be expressed as:

$$p(Y, \omega, g, \hat{f}, Z \mid X) = p(Y \mid X)\, p(\omega, g, \hat{f}, Z \mid X, Y)\,] \qquad (4)$$

However, computing the log marginal likelihood $\log p(Y \mid X)$ directly is often intractable. To address this, the model is trained by maximizing the Evidence Lower Bound (ELBO), a variational approximation of the true marginal likelihood. We denote this objective function by L, and it integrates both the data likelihood and the regularization imposed by the physics-based constraints. This ELBO-based approach was first formalized in the context of sparse Gaussian Processes by Titsias (2009).

Figure 2 shows the block diagram of the algorithmic workflow, that we have implemented to predict traffic flow and speed, for the multi-class physics-regularized Gaussian Process. The model input X is the position of the sensor and the time where the data are measured, while the output Y

are the real values for flow and speed. L is the objective function, and the execution ends when a predefined number of iterations has finished, set to 500, or the value of the objective function does not change for more than 10 iterations. We have implemented the algorithm in Tensorflow framework and to update the parameters, ADAM optimizer is used.

Figure 2 presents the block diagram of the algorithmic workflow we have implemented for traffic flow and speed prediction using the multi-class physics-regularized Gaussian Process model. The model input X consists of the sensor location and time stamp at which traffic measurements are recorded, while the output Y corresponds to the observed values for flow and speed. The training process optimizes the objective function L and terminates either when a predefined maximum number of 500 iterations is reached, or when the value of L has not changed for more than 10 consecutive iterations. The algorithm is implemented using the TensorFlow framework, and parameter updates are carried out using the ADAM optimizer.

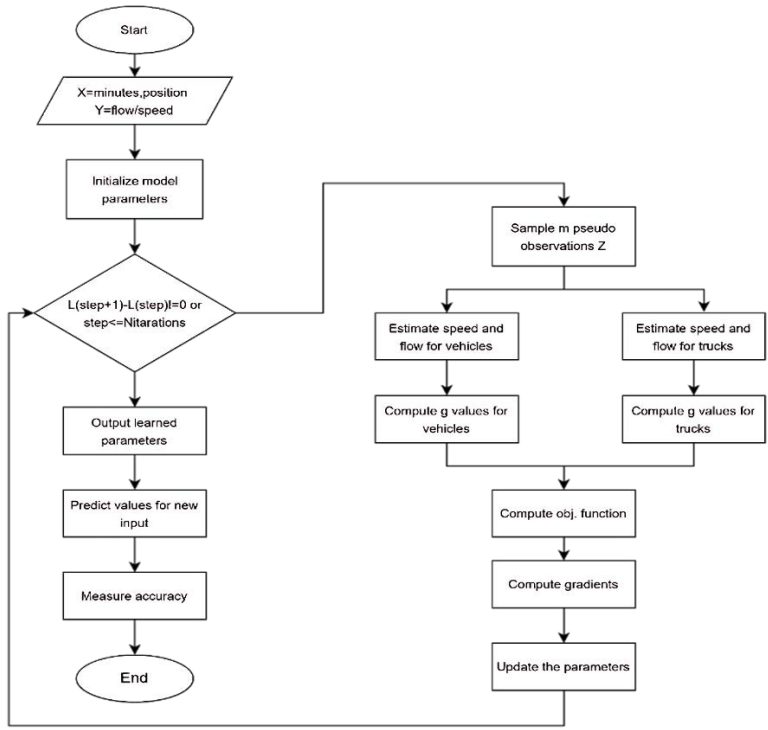


Fig. 2: Workflow of the multi-class physics-regularized Gaussian Process for traffic flow and speed prediction.

3.4.1 The physics-informed LSTM model

The physics-informed LSTM model extends the concept of incorporating physical constraints into machine learning, specifically for sequential data. Long Short-Term Memory (LSTM) networks are an advanced class of Recurrent Neural Networks (RNNs) designed to capture long-range dependencies in time-series data (Hochreiter and Schmidhuber, 1997). Unlike standard RNNs, which struggle with long-term memory due to vanishing gradients, LSTMs introduce memory cells and gating mechanisms that regulate the flow of information, enabling them to retain relevant context over extended sequences.

In contrast to the physics-regularized Gaussian Process, which enforces physical consistency through an external regularization term, the physics-informed LSTM model directly embeds the multi-class METANET equations into the network's temporal architecture. This makes it particularly well-suited for traffic flow modelling, where time-dependent dynamics play a critical role.

The model balances data-driven learning and physical consistency by incorporating a physics-based loss term into its training objective. The overall loss function consists of two components: (1) the standard mean squared error (MSE) between the LSTM predictions and observed traffic data, and (2) an additional term that penalizes discrepancies between the LSTM outputs and the physics-based estimates derived from the multi-class METANET model:

$$\mathcal{L} = \alpha \cdot \mathrm{MSE}(y, \hat{y}) + (1 - \alpha) \cdot \mathrm{MSE}\left(\hat{y}, \widehat{y'}\right) \qquad (5)$$

where $\alpha \in [0, 1]$ balances physical laws with data-driven learning. While the parameters of the multi-class METANET model are pre-calibrated using real traffic data, the parameters of the LSTM network are optimized during training using gradient-based methods, specifically Stochastic Gradient Descent (SGD) and the Adam optimizer. Figure 3 illustrates the block diagram of the algorithmic workflow implemented to predict traffic flow and speed using the multi-class physics-informed LSTM model. The input to the model is a time series of past values for a single traffic state variable (either speed or flow) across consecutive time steps. Training proceeds for a maximum of 500 iterations, or until the objective function L shows no significant improvement for more than 10 consecutive iterations. The entire algorithm is implemented in the TensorFlow framework.

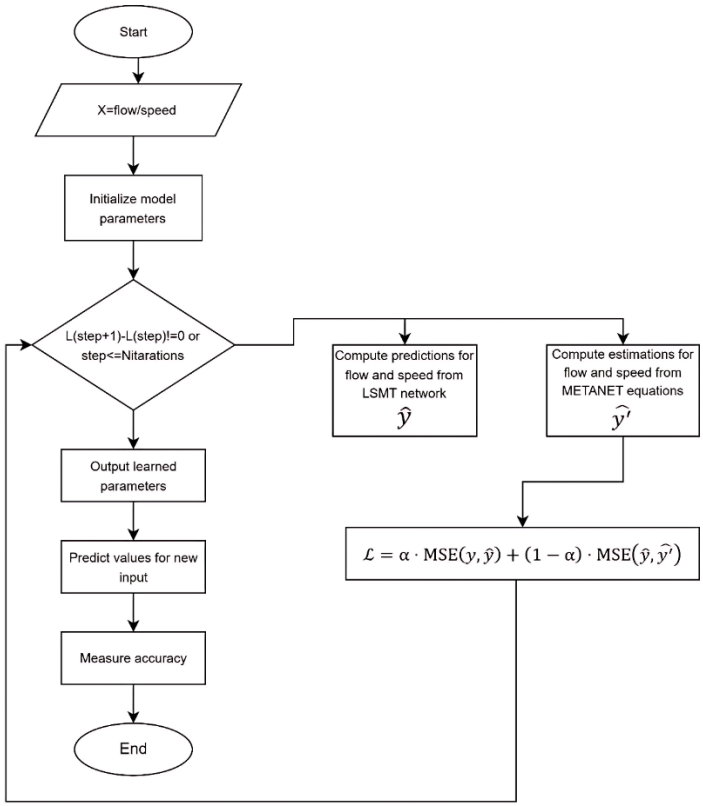

Start
X=flow/speed
Initialize model parameters
L(step+1)-L(step)!=0 or step<=Nitarations
Compute predictions for flow and speed from LSMT network
$\hat{y}$
Compute estimations for flow and speed from METANET equations
$\hat{y'}$
Output learned parameters
Predict values for new input
Measure accuracy
End
$\mathcal{L} = \alpha \cdot \mathrm{MSE}(y, \hat{y}) + (1 - \alpha) \cdot \mathrm{MSE}(\hat{y}, \hat{y'})$

Fig. 3. Workflow of the multi-class physics-informed LSTM model for traffic flow and speed prediction.

4. CASE STUDY APPLICATION

We evaluated the proposed physics-guided models using data from the Caltrans Performance Measurement System (PeMS). The selected freeway segment and the corresponding sensor locations are depicted in Figure 4. In the figure, green labels indicate mainline sensors used for training the models, blue and purple labels represent sensors that provide on-ramp and off-ramp flow measurements, and the red label marks the location where predictions are performed. Traffic data were collected at 5-minute intervals over a 5-day period from each sensor. This data collection strategy is based on findings from a prior study (Binjaku *et al*. 2025) in which we analysed the effect of training data size on the performance of a

multi-class physics-regularized Gaussian Process for traffic state prediction. That study showed that using two weeks of training data yielded the highest prediction accuracy, but at the cost of significantly increased computational time. In contrast, using 5 days of data provided an optimal trade-off between prediction accuracy and computational efficiency, and was therefore adopted in the present study.

Fig. 4: The considered freeway stretch.

Figure 5 illustrates the traffic data distribution for both cars and trucks over a five-day period at sensor S2. As shown in the plots, vehicle volumes increase significantly during peak hours and decline sharply—almost to zero—during night time. The speed distribution along the corridor similarly indicates congestion during rush hours, characterized by reduced average speeds. This traffic pattern is observed consistently across the entire study segment, confirming the temporal regularity of congestion during peak demand periods.

The Pearson correlation coefficient between the flow of vehicles and trucks is −0.06, and between their speeds is −0.09, indicating negligible linear correlation in both cases. These low correlation values suggest that vehicles and trucks exhibit distinct flow and speed patterns across the dataset. As a result, it is justifiable and appropriate to model the two vehicle classes separately, particularly in the context of multi-class traffic flow modelling.

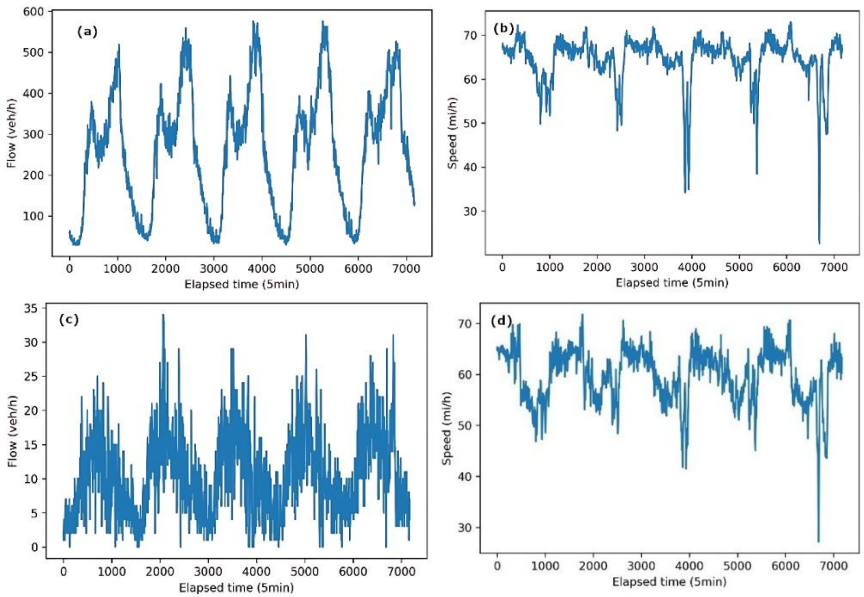


Fig.5: Flow and speed distribution during 5 days, for cars (5a and 5b) and trucks (5c and 5d).

The multi-class physics-regularized Gaussian Process (PR-GP) model was configured with two restarts and optimized using the L-BFGS-B optimizer. The multi-class physics-informed LSTM (PI-LSTM) model was implemented using three stacked LSTM layers, each comprising 50 units, followed by a Dropout layer with a dropout rate of 0.2 to mitigate overfitting. The Adam optimizer was employed for gradient-based parameter updates.

Both models were trained for 500 epochs, with a learning rate of 0.1 and a batch size of 32. Model predictions were generated for sensor S4, focusing on peak traffic hours over a single day.

The fixed parameters of the multi-class METANET model are listed in Table 1, while the remaining parameters were calibrated using the initial values provided in Table 2.

Table 1. METANET parameters with fixed value.

Parameters	**Values**
T	1/12 h
L_i	1.2 mi
λ_i	5

Table 2. METANET parameters initial values.

Parameters	Values
v_{free}	100 km/h
ρ^{cr}	59.4 veh/mi
χ	20.97 veh/mi
τ	0.005 h
α	1.5
ν	13.5 mi^2/h

Figures 6 and 7 present the predicted speed and flow for cars and trucks, respectively, using the multi-class physics-regularized Gaussian Process and the multi-class physics-informed LSTM models.

The corresponding Mean Absolute Percentage Error (MAPE) values for speed and flow predictions are reported in Tables 3 and 4, respectively. Additionally, Table 5 summarizes the training time required by each model. Training time was recorded on a standard laptop equipped with 16 GB of RAM, a 2.0 GHz quad-core CPU, and running the Windows operating system.

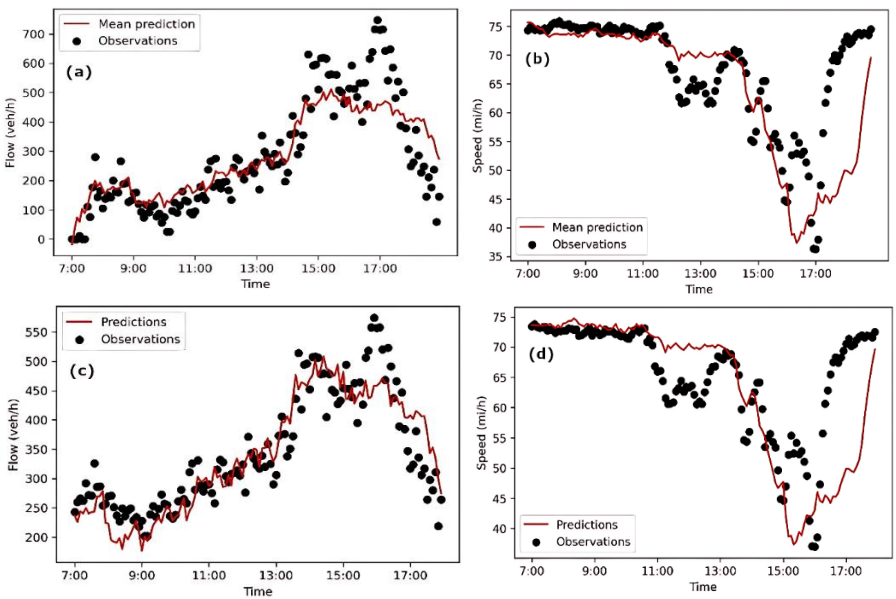


Fig. 6: Flow and speed predictions for cars, with the mono-class physics-regularized Gaussian Process (6a and 6b) and with the multi-class physics-regularized Gaussian Process (6c and 6d).

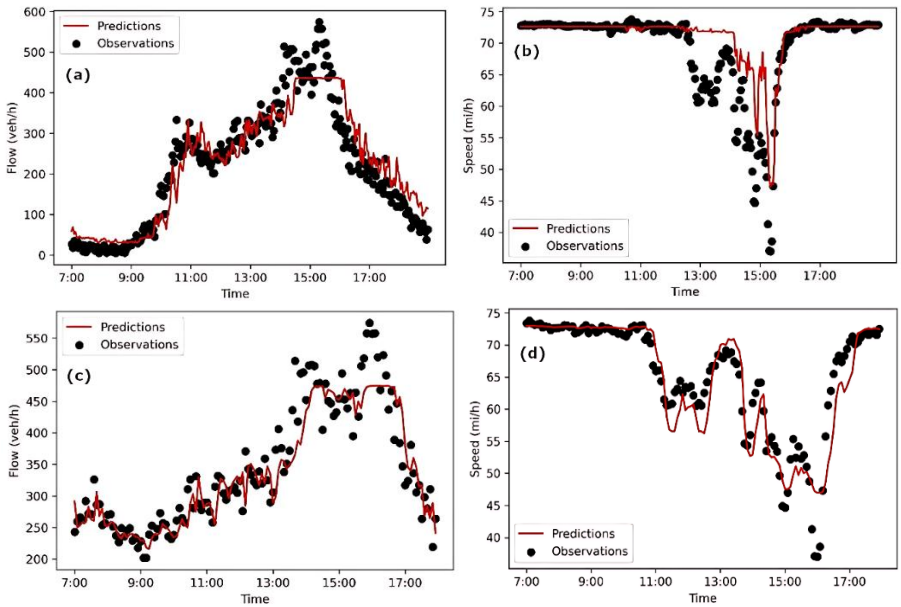


Fig. 7: Flow and speed predictions for cars, with the mono-class physics-informed LSTM (7a and 7b) and with the multi-class physics-informed LSTM (7c and 7d).

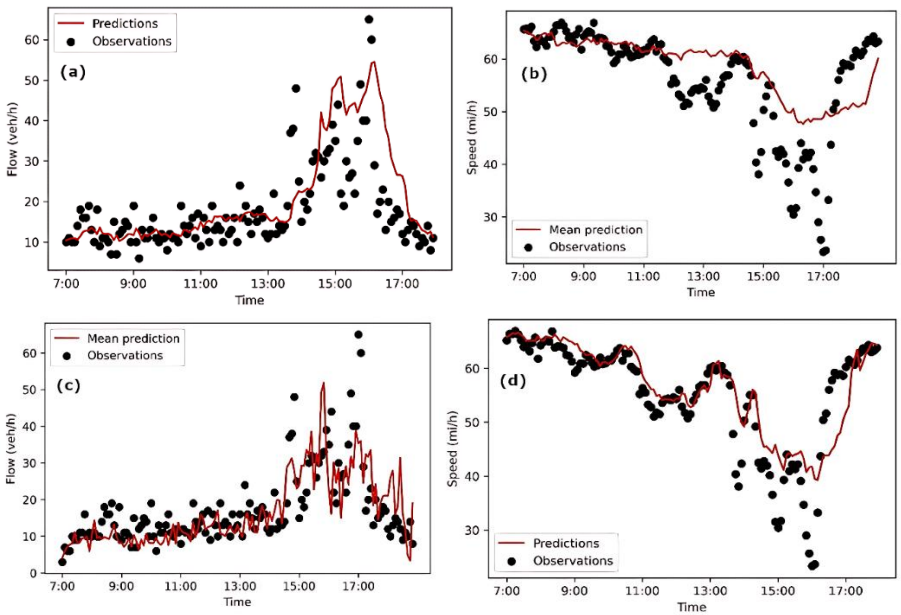


Fig. 8: Flow and speed predictions for trucks, with physics-regularized Gaussian Process (8a and 8b), and physics-informed LSTM model (8c and 8d).

Table 3. Mean Absolute Percentage Errors in flow and speed for cars, with physics-regularized Gaussian process and physics-informed LSTM models

	Traffic flow	**Traffic speed**
Mono-class physics-regularized Gaussian process	0.23	0.17
Mono-class physics-informed LSTM	0.18	0.15
Multi-class physics-regularized Gaussian Process	0.21	0.16
Multi-class physics-informed LSTM	0.17	0.11

Table 4. Mean Absolute Percentage Errors in flow and speed for trucks, with physics-regularized Gaussian process and physics-informed LSTM models

	Traffic flow	**Traffic speed**
Multi-class physics-regularized Gaussian Process	0.29	0.19
Multi-class physics-informed LSTM	0.19	0.16

Table 5. Training time

	Time [sec]
Multi-class physics-regularized Gaussian Process	18.200
Multi-class physics-informed LSTM	2109

The plots, prediction error tables, and the training time table clearly indicate that our study evaluates two key aspects: (i) the impact of transitioning from a mono-class to a multi-class modelling approach, and (ii) the trade-offs between the two model types—multi-class physics-informed LSTM and multi-class physics-regularized Gaussian Process.

First, the advantages of incorporating traffic heterogeneity are evident when contrasting mono-class and multi-class scenarios. Prediction errors are consistently higher in the mono-class setting, where vehicle types are not differentiated. For instance, the multi-class version of the physics-regularized Gaussian Process model reduces the MAPE to 0.21 for flow and 0.16 for speed, compared to 0.23 and 0.17, respectively, in the mono-class version. A similar pattern is observed with the physics-informed LSTM model: the mono-class version yields MAPE values of 0.18 (flow) and 0.15 (speed), while the multi-class version improves performance to 0.17 and 0.11, respectively. Although these reductions may seem modest in absolute terms, they are consistent and meaningful—especially in real-time traffic control scenarios where predictive precision is critical. The multi-class models enhance the network's ability to capture traffic patterns

by enabling the modelling of each vehicle class (e.g., cars and trucks) with distinct dynamics.

Second, the trade-off between training time and prediction accuracy has significant implications for model architecture and overall performance. The multi-class physics-informed LSTM model achieves the best predictive performance, particularly in the multi-class case, with error levels of 0.17 for flow and 0.11 for speed. This improvement stems from the LSTM's inherent ability to effectively capture temporal dynamics and long-term dependencies in traffic data.

In addition to its superior predictive accuracy for both flow and speed, the multi-class physics-informed LSTM model is also computationally efficient, requiring only 2,109 seconds for training. In contrast, the multi-class physics-regularized Gaussian Process (PRGP) model demands significantly more computational resources, with a training time of 18,200 seconds.

This substantial difference in training time stems from the fundamental nature of the two models. The physics-informed LSTM (PI-LSTM) relies on iterative, gradient-based optimization over multiple epochs, whereas the PRGP model depends on matrix-based computations. When physical knowledge is incorporated into the PRGP framework, the regularization scheme and the covariance structure must be adapted accordingly, increasing computational complexity. PRGP's kernel-based architecture becomes more cumbersome when integrating physical constraints. In contrast, PI-LSTM enables a more modular and flexible integration of physical knowledge through modifications to the loss function, enhancing its scalability and adaptability to diverse traffic dynamics.

5. CONCLUSIONS

In this study, we implemented and evaluated two models for traffic state prediction: the physics-informed Long Short-Term Memory (PI-LSTM) model and the physics-regularized Gaussian Process (PRGP) model, applied to both mono-class and multi-class traffic data. The models were tested using a real-world highway dataset for predicting vehicle flow and speed.

The results indicate that multi-class models consistently outperform their mono-class counterparts by achieving higher prediction accuracy through the incorporation of vehicle-specific dynamics.

Among the evaluated approaches, the multi-class PI-LSTM model demonstrated superior predictive performance while also significantly reducing training time. Its architecture is comparatively simpler to implement, requiring only modest adjustments to the loss function to incorporate physical constraints. In contrast, although the PRGP model also benefits from physics-based regularization, its implementation is more complex due to the need for modifying kernel structures and managing computationally intensive matrix operations. This leads to higher computational costs, making it less suitable for real-time applications.

Overall, the findings suggest that the multi-class physics-informed LSTM model offers the best trade-off between predictive accuracy, computational efficiency, and implementation complexity. It stands out as a practical and effective solution for real-time traffic prediction in heterogeneous traffic environments.

DECLARATIONS

Data Accessibility

All data used in this study are publicly available and have been previously published or referenced in scientific sources.

Author Contributions:

Kleona Binjaku had full access to all the data and assumes responsibility for the accuracy and integrity of the data and results. The study concept, methodological design, and model development were primarily led by Binjaku, with substantial input from Kajo Meçe and Sacone. Data acquisition, preprocessing, and model training were performed by Binjaku, under the supervision of Pasquale and Sacone. All authors contributed to the interpretation of results and formulation of the comparative framework. The initial manuscript draft was written by Binjaku. Kajo Meçe, Pasquale, and Sacone provided critical revisions and intellectual guidance throughout the writing process.

Author Approval

All authors reviewed and approved the final version of the manuscript.

Declaration of AI use:

No generative AI has been employed in writing this article.

Conflict of Interest Declaration

The authors declare no conflict of interest.

Funding

This research received no external funding.

REFERENCES:

Barreau M, Aguiar M, Liu J, Johansson KH. 2021b. Physics-informed learning for identification and state reconstruction of traffic density. *In: Proc. IEEE Conf. Decision and Control (CDC)*, 2653–2658. DOI: https://doi.org/10.1109/CDC45484.2021.9683295.

Barreau M, Liu J, Johansson KH. 2021a. Learning-based state reconstruction for a scalar hyperbolic PDE under noisy Lagrangian sensing. *In: Proc. Machine Learning Research*, 144, 34–46. https://proceedings.mlr.press/v144/barreau21a/barreau21a.pdf.

Binjaku K, Kajo Meçe E, Pasquale C, Sacone S. 2024. A comparison between Gaussian Processes and LSTM networks on implementing physics-guided machine learning models for traffic state prediction. In *The international conference: AI Horizon - A new era of creativity and innovation, 272-281. The international conference : AI Horizon - A new era of creativity and innovation : proceedings book / Akademia e Shkencave e Shqipërisë :: COBISS+*.

Binjaku K, Pasquale C. Kajo Meçe E, Sacone S. 2025. Freeway traffic modeling by physics-regularized Gaussian Processes. In *IEEE Open Journal of Intelligent Transportation Systems*, **6**, pp. 116-130. DOI: https://doi.org/10.1109/OJITS.2025.3532796.

Chavoshi K, Ferrara A, Kouvelas A. 2023. A feedback linearization approach for coordinated traffic flow management in highway systems. In *Control Engineering Practice*, **139:**105615. DOI: https://doi.org/10.1016/j.conengprac.2023.105615.

Chu Q, Li G, Zhou R, Ping Z. 2021. Traffic Flow Prediction Model Based on LSTM with Finnish Dataset. In: *Proc. International Conference on Intelligent Computing and Signal Processing*, 389–392. DOI: https://doi.org/10.1109/ICSP51882.2021.9408888.

Das T, Tanvir S. 2024. Emissions aware car following Model: A physics informed LSTM application. In *Proc. IEEE Forum for Innovative Sustainable Transportation Systems (FISTS)*, 1–5. DOI: https://doi.org/10.1109/FISTS60717.2024.10485538

Di X, Shi R, Mo Z, Fu Y. 2023. Physics-Informed Deep Learning for Traffic State Estimation: A Survey and the Outlook. *In: Algorithms*, **16(6):** 305. https://www.mdpi.com/1999-4893/16/6/305.
DOI**:** https://doi.org/10.1109/CVIDL51233.2020.00-70.
Fafoutellis P, Vlahogianni EI. 2025. A theory-informed multivariate causal framework for trustworthy short-term urban traffic forecasting. *In: Transportation Research Part C: Emerging Technologies*, **170 (6):**104945. Doi: https://doi.org/10.1016/j.trc.2024.104945.
Geng M, Li J, Xia Y, Chen X. (M.). 2023. A physics-informed Transformer model for vehicle trajectory prediction on highways. *In: Transportation Research Part C: Emerging Technologies*, **154:** 104272. DOI: https://doi.org/10.1016/j.trc.2023.104272.
Han Y, Wang M, Li L, Roncoli C, Gao J, Liu P. 2022. A physics-informed reinforcement learning-based strategy for local and coordinated ramp metering. *In: Transportation Research Part C: Emerging Technologies*, **137:**103584. DOI: https://doi.org/10.1016/j.trc.2022.103584.
Hegyi A, De Schutter B, Hellendoorn J. 2005. Optimal coordination of variable speed limits to suppress shock waves. In *IEEE Transactions on Intelligent Transportation Systems*, **6 (1):** 102–112. DOI: https://doi.org/10.1109/TITS.2004.842408.
Hochreiter S, Schmidhuber J. 1997. Long short-term memory. In *Neural Computing*, **9 (8):** 1735–1780. DOI: https://doi.org/10.1162/neco.1997.9.8.1735.
https://doi.org/10.1016/j.commtr.2024.100142.
Huang AJ, Agarwal S. 2020. Physics informed deep learning for traffic state estimation. *In: Proc. 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–6. DOI: https://doi.org/10.1109/ITSC45102.2020.9294236.
Huang AJ, Agarwal S. 2021. Physics Informed Deep Learning: Applications to traffic state estimation in connected vehicle environment. In: Proc. Transportation Research Board 100th Annual Meeting.
https://www.researchgate.net/publication/349255408_Physics_Informed_Deep_Learning_Applications_to_Traffic_State_Estimation_in_Connected_Vehicle_Environment.
Huang AJ, Agarwal S. 2022. Physics-informed deep learning for traffic state estimation: Illustrations with LWR and CTM Models. *In: IEEE*

Open Journal of Intelligent Transportation Systems, **3:** 503–518. DOI: https://doi.org/10.1109/OJITS.2022.3182925.

Huang AJ, Agarwal S. 2023a. Physics informed deep learning: Applications in transportation. https://doi.org/10.48550/arXiv.2302.12336.

Huang AJ, Agarwal S. 2023b. On the limitations of physics-informed deep learning: illustrations using first-order hyperbolic conservation law-based traffic flow models. *In: IEEE Open Journal of Intelligent Transportation Systems*, **4**: 279–293. DOI: https://doi.org/10.48550/arXiv.2302.12337.

Huang AJ, Biswas A, Agarwal S. 2024. Incorporating nonlocal traffic flow model in physics-informed neural networks. *In: IEEE Transactions on Intelligent Transportation Systems*, **25(11):** 16249–16258. DOI: https://doi.org/10.48550/arXiv.2308.11818.

Ikotun AM, Ezugwu AE, Abualigah L, Abuhaija B, Heming J. 2023. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. In *Information Science* **622**, 178–210. DOI: https://doi.org/10.1016/j.ins.2022.11.139.

Kocijan J, Přikryl J. 2010. Soft sensor for faulty measurements detection and reconstruction in urban traffic. In: *Proc. IEEE Mediterranean Electrotechnical Conference*, 172–177. Available online 15.07. 2025 https://library.utia.cas.cz/separaty/2010/AS/prikryl-soft%20sensor%20for%20faulty%20measurements%20detection%20and%20reconstruction%20in%20urban%20traffic.pdf.

Lei N, Zhang H, Li R, Yu J, Wang H, Wang Z. 2024. Physics-informed data-driven modeling approach for commuting-oriented hybrid powertrain optimization. *In: Energy Conversion and Management*, **299**. DOI: https://doi.org/10.1016/j.enconman.2023.117814.

Lian R, Wang X. 2024. GCN-Bi-LSTM: A neural network method based on spatiotemporal features for short-term traffic flow prediction. In*: Proc. International Conference on Data-driven Optimization of Complex Systems*, 613–618. DOI: https://doi.org/10.1109/DOCS63458.2024.10704288.

Liu D, Hui S, Li L, Liu Z, Zhang Z. 2020. A method for short-term traffic flow forecasting based on GCN-LSTM. In: *Proc. International Conference on Computer Vision, Image and Deep Learning*, 364–368.

Liu J, Jiang R, Zhao J, Shen W. 2023. A quantile-regression physics-informed deep learning for car-following model. *In: Transportation*

Research Part C: Emerging Technologies, **154(2):** 104275. DOI: https://doi.org/10.1016/j.trc.2023.104275.

Liyong W, Vateekul P. 2019. Traffic prediction using attentional spatial-temporal deep learning with accident embedding. In: *Proc. International Conference on Computational Intelligence and Applications*, 98–103. DOI: https://doi.org/10.1109/ICCIA.2019.00026.

Long K, Shi X, Li X. 2024. Physics-informed neural network for cross-dynamics vehicle trajectory stitching. In *Transportation Research Part E: Logistics and Transportation Review*, **192**: 103799. DOI: https://doi.org/10.1016/j.tre.2024.103799.

Mo Z, Shi R, Di X. 2021. A physics-informed deep learning paradigm for car-following models. *In: Transportation Research Part C: Emerging Technologies*, 130, 103240. DOI: https://doi.org/10.1016/j.trc.2021.103240.

Papageorgiou M, Blosseville JM, Hadj-Salem H. 1989. Macroscopic modeling of traffic Flow on the Boulevard Périphérique in Paris. In *Transportation Research Part B*, **23 (1)**: 29–47. DOI: https://doi.org/10.1016/0191-2615(89)90021-0.

Pasquale C, Papamichail I, Roncoli C, Sacone S, Siri S, Papageorgiou M. 2015. Two-class freeway traffic regulation to reduce congestion and emissions via nonlinear optimal control. In *Transportation Research C*, **55**: 85–99. DOI: https://doi.org/10.1016/j.trc.2015.01.013.

Pasquale C, Sacone S, Siri S. 2014. Two-class emission traffic control for freeway systems. In *IFAC Proceedings* **473 (3)**: 936–941. DOI: https://doi.org/10.3182/20140824-6-ZA-1003.01995.

Pereira M, Lang A, Kulcsár B. 2022. Short-term traffic prediction using physics-aware neural networks. *In: Transportation Research Part C: Emerging Technologies*, 142:103772. DOI: https://doi.org/10.1016/j.trc.2022.103772.

Rempe F, Loder A, Bogenberger K. 2021. Estimating motorway traffic states with data fusion and physics-informed deep learning. *In: Proc. IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2208–2214. DOI: https://doi.org/10.1109/ITSC48978.2021.9565096.

Shahri PK, HomChaudhuri B, Ghaffari A, Ghasemi AH. 2022. Designing traffic management strategies for a heterogeneous traffic

network. In: *IFAC-PapersOnLine* **55(37):** 694–699. DOI: https://doi.org/10.1016/j.ifacol.2022.11.263.

Sheng Z, Huang Z, Chen S. 2024. Traffic expertise meets residual RL: Knowledge-informed model-based residual reinforcement learning for CAV trajectory control. *In: Communications in Transportation Research*, **4:** 100142. DOI: https://doi.org/10.1016/j.commtr.2024.100142.

Sheng Z, Huang Z, Chen S. 2024. Traffic expertise meets residual RL: Knowledge-informed model-based residual reinforcement learning for CAV trajectory control. *In: Communications in Transportation Research*, **4:**100142. DOI:

Shi R, Mo Z, Di X. 2021. Physics-informed deep learning for traffic state estimation: A hybrid paradigm informed by second-order traffic models. *In: Proc. AAAI Conference on Artificial Intelligence*, 35(1): 540-547. DOI: https://doi.org/10.1609/aaai.v35i1.16132.

Shi R, Mo Z, Huang K, Di X, Du Q. 2022. A Physics-Informed Deep Learning Paradigm for Traffic State and Fundamental Diagram Estimation. *In: IEEE Transactions on Intelligent Transportation Systems*, **23(8):** 11688–11698. DOI: https://doi.org/10.1109/TITS.2021.3106259.

Su H, Zhang H. 2017. Online Updating for Gaussian Process Learning. In *International Conference on Computational Science and Computational Intelligence (CSCI)*, 180–183. DOI. https://doi.org/10.1109/CSCI.2017.28.

Sun S, Xu X. 2011. Variational inference for infinite mixtures of Gaussian processes with applications to traffic flow prediction. In *IEEE Transactions on Intelligent Transportation Systems*, **12 (2):** 466–475. DOI: https://doi.org/10.1109/TITS.2010.2093575.

Tian Y, Zhang K, Li J, Lin X, Yang B. 2018. LSTM-based traffic flow prediction with missing data. In *Neurocomputing*, **318 (5):** 297–305. DOI**:** https://doi.org/10.1016/j.neucom.2018.08.067.

Tischmann P, Baumann R, Stockem Novo A. 2024. Physics-informed deep learning for motion prediction in autonomous driving. In: *Proc. AmEC – Automotive meets Electronics & Control; 14th GMM Symposium*, Available online: 7–12. https://nuraxys.com/wp-content/uploads/vde-ame2023-download-data.pdf.

Titsias MK. 2009. *Variational Learning of Inducing Variables in Sparse Gaussian Processes*. In AISTATS. DOI: http://proceedings.mlr.press/v5/titsias09a/titsias09a.pdf.

Usama M, Ma R, Hart J, Wojcik M. 2022. Physics-informed neural networks (PINNs)-based traffic state estimation: An application to traffic network. *In: Algorithms*, **15 (12):** 447. https://doi.org/10.3390/a15120447.

Wang S, Yan C, Shao Y. 2023. Road traffic accident prediction model based on J-LSTM+attention mechanism. *In: Proc. Int. Conf. Artificial Intelligence and Big Data (ICAIBD)*, 635–638. DOI: https://doi.org/10.1109/ICAIBD57115.2023.10206203.

Xie H, Hu D, Song K. 2021. An iterative optimization algorithm for vehicle speed prediction considering driving style and historical data effects. In: *IEEE Chinese Control Conference*, 6094–6100. DOI: https://doi.org/10.23919/CCC52363.2021.9550541.

Zhan H, Gomes G, Li XS, Madduri K, Sim A, Wu K. 2018. Consensus ensemble system for traffic flow prediction. In *IEEE Transactions on Intelligent Transportation Systems*, **19 (12):** 3903–3914. DOI: https://ieeexplore.ieee.org/document/8306460/.

Zhao C, Yu H. 2022. Integrating PDE observer with deep learning for traffic state estimation. *In: Proc. IEEE Int. Conf. Intelligent Transportation Systems (ITSC)*, 1964–1969. https://amslab.org/wp-content/uploads/2024/10/integrating-pde-observer-with-deep-learning-for-traefac80ic-state-estimation.pdf.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

LEARNPLC: A JAVAFX-BASED EDUCATIONAL PLATFORM FOR LADDER LOGIC PROGRAMMING WITH REAL-TIME CIRCUIT SIMULATION

Ina PAPADHOPULLI and Martin ÇARO
Department of Computer Engineering, Faculty of Information Technology, Polytechnic University of Tirana, Albania
Author for correspondence: ipapadhopulli@fti.edu.al
IP, 0000-0003-2040-6860; MC, 0009-0003-7513-6797
https://doi.org/10.67537/ajnts3001005

ABSTRACT

Efficient education in programmable logic controller (PLC) programming requires tools that integrate theoretical knowledge with practical applications. This paper presents **LearnPLC**, a comprehensive educational platform developed in JavaFX designed to provide an intuitive environment for ladder logic programming and real-time circuit simulation. The platform employs a directed graph model for circuit representation and integrates real-time signal propagation algorithms to replicate the behaviour of industrial PLC. It supports standard PLC components, including contacts, coils, timers (TON/TOF), and counters (CTU/CTD), while delivering immediate visual feedback during the simulation. By combining technical precision with a user-friendly interface, LearnPLC addresses the common challenges in PLC education and enhances hands-on learning. Results proved the platform's effectiveness in teaching automation concepts through realistic and interactive circuit behaviour.

1. INTRODUCTION

Programmable Logic Controller (PLC) programming education is crucial for preparing skilled automation engineers and addressing the growing skills gap in industrial automation. However, many educational institutions face challenges in providing effective PLC training due to the limited availability of accessible, modern tools that deliver realistic learning experiences without the complexity typically associated with industrial-grade solutions (https://www.siemens.com/tia-portal).

To develop practical skills, students need hands-on experience in ladder logic programming, circuit construction and real-time simulation. However, educational environments demand specialized tools that prioritize learning effectiveness, immediate feedback, and progressive skill

development rather than extensive feature sets designed for industrial production environments (Petruzella, 2016).

LearnPLC addresses these challenges through two key innovations:

1. **Directed- graph-based circuit representation:** Ladder logic circuits are modelled as graphs, with nodes representing the PLC elements and edges representing the electrical connections. This approach enables efficient simulation algorithms while maintaining consistency with the established electrical circuit principles.
2. **Real-time simulation engine:** The platform incorporates a simulation engine that provides immediate visual feedback, including the visualization of signal propagation. This feature helps students understand the dynamic behaviour of PLC circuits without the need for physical circuit hardware.

The platform supports standard PLC elements, such as contacts, coils, timers, and counters, enabling an accurate representation of fundamental automation concepts. The main educational objectives addressed include circuit construction, real-time simulation, and visual state monitoring, which are critical components of effective automation learning.

To evaluate the platform's effectiveness, we analyse it with respect to three key factors:

- **Functional completeness** relative to educational requirements
- **User experience and accessibility** for educational contexts
- **Implementation architecture and educational design**

The remainder of the paper is structured as follows: Section 2 provides an overview of related work in PLC education, existing simulation platforms, and the theoretical foundations underlying our approach. Section 3 describes the proposed platform architecture and its implementation details. Section 4 presents the technical evaluation and analysis of the study. Section 5 discusses the educational effectiveness and benefits of the proposed method. Finally, Section 6 concludes with the key findings and outlines the directions for future research.

2. RELATED WORK

The literature on PLC education and simulation platforms highlights several approaches to address the challenges of automation education. In this study, we examined existing educational tools, research-oriented solutions, and the theoretical foundations that inform our platform design.

2.1 Educational PLC Simulation Platforms and the Need for a New System

Current PLC programming education faces persistent challenges that existing simulation platforms do not adequately address. A review of the available educational tools revealed recurring limitations, which motivated the development of LearnPLC.

Limitations of Existing Platforms

PLCLogix simplifies programming interfaces but lacks the real-time visual feedback necessary for understanding dynamic circuit behaviour (PLCLogix, https://www.plclogix.com). Factory I/O provides advanced 3D visualization but introduces unnecessary complexity that can overwhelm beginners (Vargas *et al.* 2023). LogixPro and Automation Studio offer sophisticated simulation features but require costly licenses and complex setup procedures to use. Commercial industrial platforms, such as Siemens TIA Portal and Schneider Unity Pro, deliver authentic professional experiences but pose substantial barriers owing to their high costs, demanding installation requirements, and extensive feature sets (Siemens TIA Portal, https://www.siemens.com/tia-portal). Open-source solutions, such as OpenPLC, demonstrate feasibility but lack educationally tailored features and often require significant technical expertise (Alves *et al.* 2014).

Critical Educational Gaps

1. **Accessibility barriers:** High licensing costs and complex installation procedures limit widespread adoption of these technologies.
2. **Limited real-time feedback:** Many platforms do not provide immediate visualization of circuit behaviour.
3. **Inappropriate interface design:** Interfaces are often developed for professionals rather than tailored to students learning needs.
4. **Insufficient progressive learning:** Most platforms present the full range of features at once without supporting gradual skill development.
5. **Lack of educational features:** Industrial completeness is frequently prioritized over pedagogical effectiveness.

Educational Requirements

Effective PLC education requires i) immediate visual feedback to clarify cause-and-effect relationships, ii) progressive complexity management from basic to advanced concepts, iii) safe environments for experimentation, iv) accessible platforms with minimal setup and free availability, and v) a student-focused interface design that prioritizes learning effectiveness. These requirements underscore the need for LearnPLC, a purpose-built educational platform designed to combine technical accuracy and industrial relevance with accessibility, usability, and robust pedagogical support.

2.2 Technical Foundations and Software Architecture

Modern educational software development benefits from established software-engineering practices. The Model-View-Controller (MVC) architectural pattern supports the creation of maintainable and extensible educational platforms (https://docs.oracle.com/javafx/). JavaFX provides a robust foundation for developing rich cross-platform educational applications with sophisticated and responsive user interfaces.

Graph-based algorithms provide efficient foundations for circuit simulations and analysis. The challenge lies in adapting these theoretical foundations into practical educational tools that balance computational efficiency and pedagogical effectiveness (Cormen *et al*. 2022).

Research has further demonstrated the educational value of innovative approaches to PLC instruction. Pavković and Živanović (2011) highlighted the effectiveness of virtual environments for PLC programming education, while academic initiatives, such as OpenPLC (Alves *et al*. 2014), have shown the feasibility of creating comprehensive PLC solutions tailored to educational needs. These projects underscore the importance of open, accessible approaches to automation education while identifying opportunities for the development of specialized educational tools.

2.3 Software Engineering Approaches to Educational Tools

Modern educational software increasingly integrates user-centered design principles with contemporary software engineering. The use of Model–View–Controller (MVC) architectural patterns enable the development of maintainable, extensible, and modular learning platforms (https://docs.oracle.com/javafx/). JavaFX provides a robust foundation for

building rich, cross-platform educational applications with sophisticated and responsive user interfaces (Cormen *et al*. 2022).

Incorporating real-time simulation capabilities into educational platforms requires careful architectural design, with attention to performance considerations, modularity, and scalability. These considerations ensure system responsiveness while maintaining pedagogical effectiveness.

3. PROPOSED EDUCATIONAL PLATFORM ARCHITECTURE

This section introduces the LearnPLC educational platform, outlining its system architecture, circuit representation model, and real-time simulation engine. The platform integrates established software engineering principles with domain-specific requirements for PLC education to create a robust and pedagogically effective learning environment.

3.1 System Architecture Design

LearnPLC employs a modular architecture that separates concerns while maintaining a strong educational focus. The system is organized into several core functional modules, each responsible for a distinct aspect of the platform's operation. This modular design not only enhances maintainability and scalability but also facilitates targeted improvements to individual components without compromising the overall system.

Core Application Modules

User Interface Module

The LearnPLCApp functions as the main application controller, coordinating user interactions and managing the overall application state. The *StartupMenu* provides the initial user experience and project management capabilities, while the *ThemeManager* supports visual customization to accommodate different learning preferences.

Circuit Management Module

At the core of the platform is a directed-graph-based circuit model, implemented through a set of interconnected classes:

- ***CircuitDigraph*** – Implements the directed graph representation of ladder logic circuits.

- ***CircuitNode*** – Abstract base class for all circuit elements with specialized implementations.
- ***CircuitEdge*** – Manages connections between circuit elements and tracks signal states.
- ***ConnectionManager*** – Validates circuit topology and ensures connection integrity.

Element Library Module

Provides a comprehensive set of PLC element implementations:

- ***ElementNode*** – Input elements (contacts) with user-controllable states.
- ***OutputNode, SetOutputNode, ResetOutputNode*** – Various output element types.
- ***TimerNode*** – Implements ON-delay and OFF-delay timing with real-time execution.
- ***CounterNode*** – Provides up/down counting functionality with preset values.
- ***ConnectorNode*** – Junction points for parallel circuit paths.
- ***SourceNode*** – Represents the power source for circuit initiation.

Functional Support Module

Contains specialized components that define core application functionality:

- ***ElementManager*** – Enables drag-and-drop functionality and organizes the component palette.
- ***TableManager*** – Provides real-time monitoring and state display for all circuit elements.
- ***UndoRedoManager*** – Manages system states to enable safe experimentation.
- ***UniqueIDGenerator*** – Ensures proper element identification and referencing.

Project Management Module

Supports educational workflow requirements:

- ***FileManager*** – JSON-based project serialization optimized for sharing and collaboration.
- ***ApplicationSnapshot*** – Manages state persistence for undo/redo functionality and session recovery.

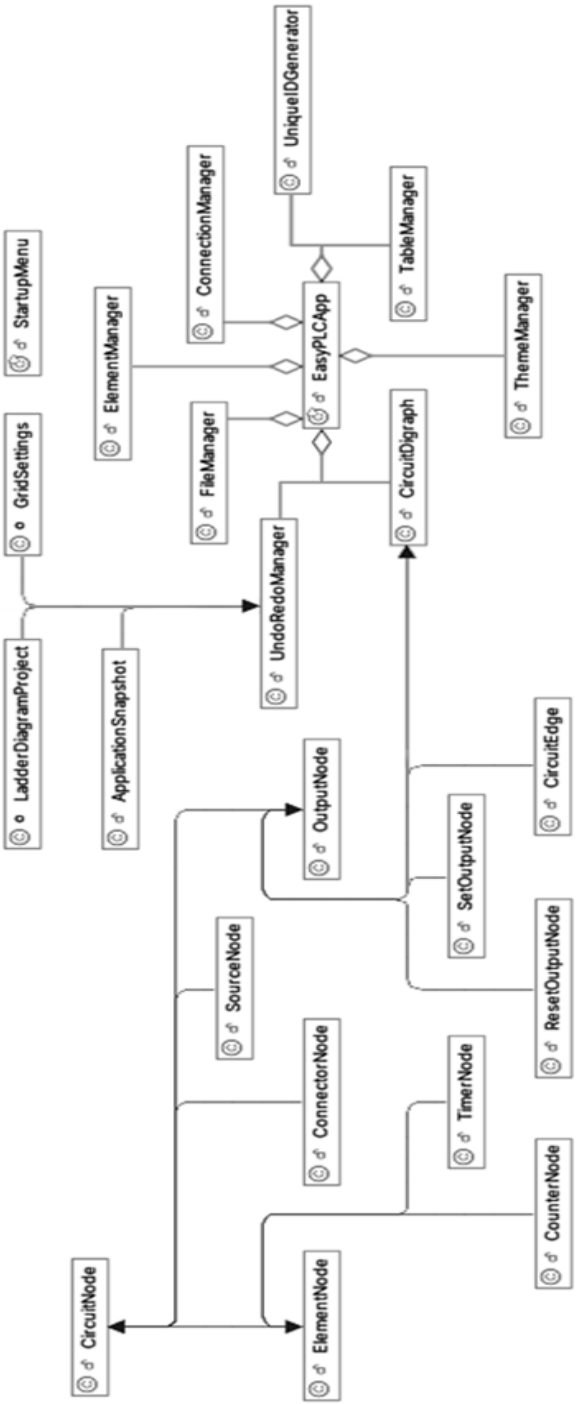


Fig. 1: Modular architecture of LearnPLCApp, illustrating user interface, circuit management, element library, functional support, and project management modules.

3.2 Circuit Representation Model

LearnPLC models ladder logic circuits as directed graphs, where nodes correspond to PLC elements and edges represent electrical connections. This representation supports efficient simulation algorithms, aligns naturally with electrical circuit principles, and enables straightforward validation of circuit integrity.

The ***CircuitNode*** abstract class serves as the foundation for all circuit elements, with specialized subclasses providing distinct different functionalities:

- ***ElementNode*:** Input elements (contacts) with user-controllable states.
- ***OutputNode*:** Output elements (coils) that respond to input signals.
- ***CounterNode*:** Supports up/down counting operations with preset values.
- ***TimerNode*:** Implements ON-delay and OFF-delay timing functions.
- ***ConnectorNode*:** Represents junction points for parallel circuit paths.

The ***CircuitEdge*** class models the electrical connections between elements, maintaining signal state and visual properties to ensure efficient traversal while preserving electrical characteristics.

3.3 Real-Time Simulation Engine

The simulation engine provides a sophisticated yet educationally accessible approach to modeling circuit behavior. It employs a modified breadth-first search (BFS) algorithm for signal propagation, ensuring performance consistent with industrial PLC operations while maintaining real-time responsiveness suitable for educational environments.

Core Simulation Architecture

The engine operates in discrete cycles that mirror industrial PLC scan processes. Each simulation cycle follows a structured sequence:

1. **Initialization Phase:** Resets visual indicators and prepares the circuit for execution.
2. **Reset Signal Processing:** Handles reset commands with assigned priority.
3. **Element State Evaluation:** Updates the internal states of timers and counters.

4. **Signal Propagation:** Traverses the circuit graph according to logical rules.
5. **Visual Update Phase:** Updates all visual indicators in real-time.

Advanced Simulation Capabilities

The engine supports a range of advanced behaviors, including:

- Priority handling for reset signals.
- Edge detection for counter inputs.
- Timer threading for background execution.
- Efficient signal propagation with logical evaluation.
- Real-time visual feedback updates for immediate comprehension.

Educational Benefits

This simulation approach enables students to observe authentic PLC behaviour without the need for physical hardware. Immediate visual feedback reinforces cause-and-effect understanding, while error visualization accelerates learning. Furthermore, adaptive complexity management ensures that the interface remains intuitive as students' progress toward mastering advanced timing and control interactions.

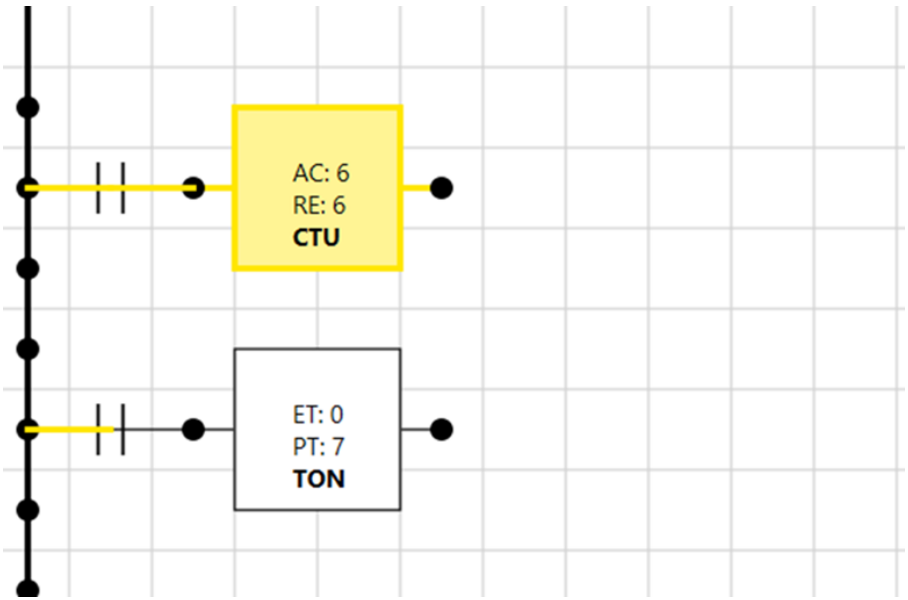


Fig. 2: Example of a simulation in action.

3.4 Advanced Educational Features

LearnPLC integrates several advanced features aimed at enhancing the educational experience:

• **Immediate Visual Feedback:** Circuit connections dynamically change color in real time to indicate signal flow, enabling students to visualize circuit behavior. Active elements are highlighted to provide a clear understanding of system operation.

• **Error Prevention:** The platform automatically validates circuit connections and element configurations to prevent common beginner errors. Invalid connections are rejected, and informative messages guide students toward correct implementation.

• **Progressive Complexity:** The component palette is organized by functional complexity, allowing instructors to introduce concepts progressively. Basic contacts and coils offer introductory-level functionality, while timers and counters support the development of more advanced automation scenarios.

• **State Monitoring:** A comprehensive monitoring panel displays the current state of all circuit elements, including input values, output states, timer progress, and counter accumulation. This real-time feedback helps students correlate programmed logic with observable system behavior.

4. IMPLEMENTATION AND TECHNICAL DETAILS

The LearnPLC educational platform was implemented and evaluated to demonstrate both its technical functionality and pedagogical effectiveness. The system was developed in Java 11 using JavaFX 17, ensuring full cross-platform compatibility across Windows, macOS, and Linux educational environments. Following established software engineering principles, a modular architecture was adopted to support current educational use while enabling future scalability and extensibility.

4.1 Development Environment and Implementation Framework

The technology stack was carefully selected to meet educational objectives while ensuring long-term maintainability and scalability. **Java 11** provides robust, cross-platform compatibility essential for heterogeneous educational computing environments, whereas **JavaFX 17** enables the development of modern, responsive user interfaces with

advanced visual effects and real-time updates. **JSON serialization** is employed to generate human-readable project files, facilitating file sharing and collaboration among students and instructors. **Multithreading** is implemented to handle background processes such as timers and real-time simulation, maintaining interface responsiveness even under complex circuit operations.

The modular design supports the independent development, testing, and maintenance of components while preserving the educational focus. Each user interface component and logical element is encapsulated as a separate module, allowing for straightforward debugging and future feature expansion. Educational design principles guide the implementation, prioritizing user experience and instructional effectiveness over raw computational performance.

Additional features designed specifically for educational environments include:

- **Classroom Mode**: Simplified interfaces optimized for live demonstrations and group activities.
- **Accessibility Support**: Compatibility with screen readers and full keyboard navigation to ensure inclusive learning.
- **Multilingual Framework**: Built-in support for internationalization to accommodate diverse educational contexts.
- **Network Integration**: Potential for linkage with learning management systems and collaborative educational platforms.

4.2 Circuit Representation and Algorithm Implementation

The core technical innovation of **LearnPLC** lies in its representation of ladder logic circuits as **directed graphs**, where nodes correspond to PLC elements and edges represent electrical connections. This design enables efficient simulation algorithms while preserving the integrity of fundamental electrical circuit principles.

The **CircuitDigraph** class functions as the central component, implementing graph traversal algorithms to manage signal propagation. The simulation engine utilizes a **modified breadth-first search (BFS)** algorithm that processes signals in discrete cycles, accurately reproducing the **scan cycle behavior** of industrial PLCs.

The signal propagation process follows a structured sequence:

1. **Initialization:** Each simulation cycle begins with the reset of edge colors and the prioritized handling of reset signals.

2. **Element Processing:** Timers and counters are updated according to the current input conditions.
3. **Signal Propagation:** A modified BFS algorithm propagates signals through the circuit, taking into account the electrical properties of each element type.
4. **Visual Updates:** Connection states and element indicators are refreshed in real-time, providing immediate feedback to the user.

```
public void simulateSignal() {
     // Initialize simulation cycle
     startNewSimulationCycle();
     resetEdgeColors();
    // Process reset signals first
     processResetSignals();

    // Handle timers and counters
     processTimersAndCounters();

    // Propagate signals through the circuit
     Queue<CircuitNode> queue = new LinkedList<>();
     addSourceNodesToQueue(queue);

    while (!queue.isEmpty()) {
         CircuitNode current = queue.poll();
         if (current.processSignal()) {
             propagateToConnectedNodes(current, queue);
         }
     }

    // Update visual feedback
     updateConnectionVisuals();
     updateElementDisplays();
 }
```

The simulation algorithm effectively handles complex scenarios, including priority-based reset signal processing, edge detection for counter inputs, background timer threading, and real-time visual updates, thereby ensuring an accurate and responsive circuit simulation environment ().

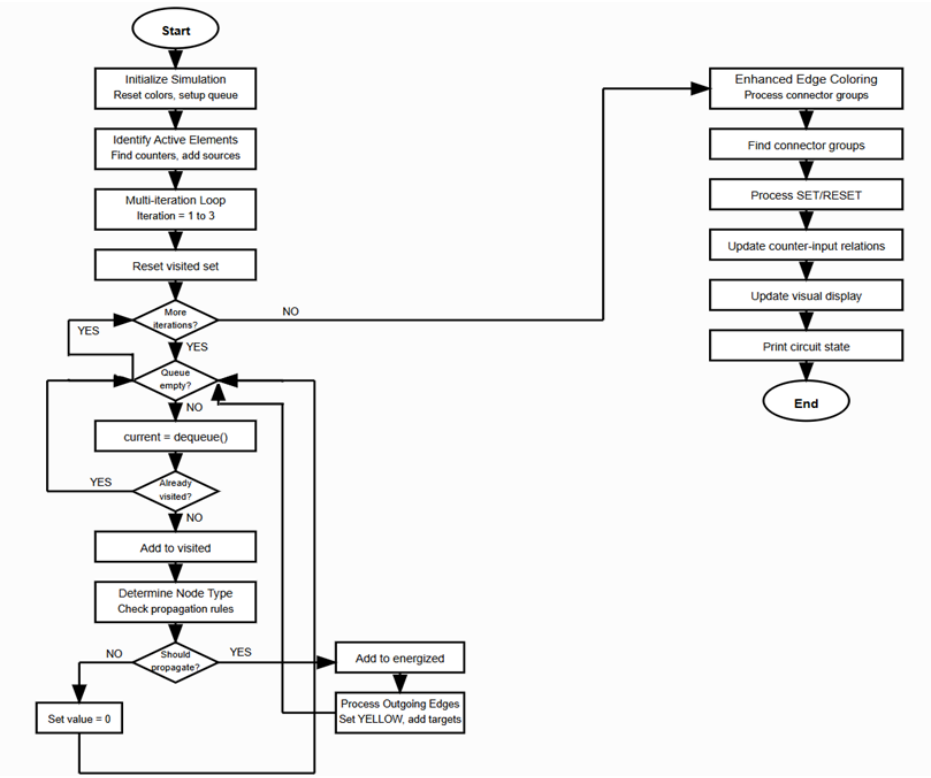


Fig. 3: Control flow of the signal simulation algorithm in LearnPLC, illustrating signal propagation, timer/counter handling, and the real-time update cycle.

4.2.1 Directed Graph Advantages in Circuit Representation

The directed graph representation offers several key advantages that make it particularly effective for educational PLC simulation:

- **Natural Circuit Modeling:** The graph structure maps directly onto ladder logic topology, where nodes represent PLC elements (contacts, coils, timers) and directed edges represent electrical connections with defined signal flow. This approach establishes an intuitive correspondence between the visual circuit layout and the underlying data model, thereby enhancing readability and conceptual understanding.
- **Efficient Simulation Algorithms:** Graph traversal techniques—particularly the modified breadth-first search (BFS) algorithm—enable systematic signal propagation throughout the circuit. The

directed nature of edges ensures that signals follow correct electrical flow patterns, moving from input elements through intermediate logic to output elements. This behavior closely mirrors that of industrial PLC systems.

- **Source Traceability and Debugging:** The directed graph structure supports efficient backward traversal for troubleshooting and educational analysis. When circuit elements are modified or removed, the system can trace connections back to their source nodes, maintaining complete circuit integrity and clearly illustrating signal dependencies. This bidirectional navigation capability is essential for interactive circuit editing and for providing students with meaningful feedback on signal origins and logical flow.

Educational Benefits: The graph abstraction allows students to focus on logical relationships rather than low-level implementation details. By working with familiar ladder logic symbols, learners develop an intuitive understanding of automation principles, while the platform's underlying graph algorithms automatically handle signal timing, state persistence, and interactions between parallel paths.

Validation Capabilities: Graph-based algorithms enable the automatic detection of common circuit configuration errors, such as unconnected elements, invalid connections, and incomplete circuit paths. This real-time validation provides immediate feedback, allowing students to identify and correct mistakes as they occur. Such instant reinforcement promotes proper circuit construction techniques and significantly enhances learning outcomes.

Extensibility: The modular system design allows new PLC element types to be introduced easily by implementing the *CircuitNode* interface. These newly added elements automatically inherit the platform's graph traversal and simulation capabilities without requiring changes to the core algorithms. This approach supports the gradual integration of advanced automation concepts and ensures the platform's long-term scalability and maintainability.

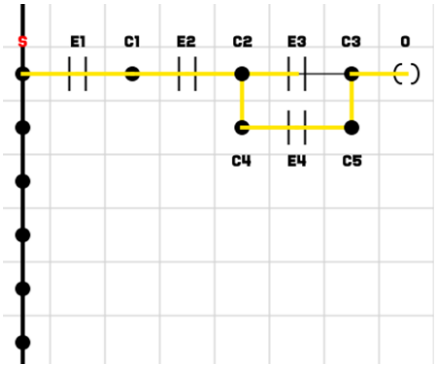


Fig. 4: Circuit example.

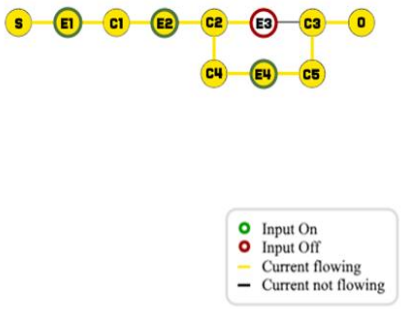


Fig. 5: Digraph implementation of the circuit.

4.3 Educational Feature Implementation and User Interface Design

The LearnPLC user interface is developed in JavaFX and follows established educational usability principles. Its three-panel layout provides a clear and intuitive workspace comprising:

- **Component Palette** for drag-and-drop circuit construction
- **Central Grid Area** for circuit development
- **Real-Time Monitoring Panel** for observing system behavior

Component Palette Design: The component palette enables intuitive circuit construction without requiring programming syntax. PLC elements are organized by functional complexity to support progressive learning. Grid-based positioning ensures orderly layouts, helping students maintain visually structured circuits as their projects increase in complexity.

Visual Feedback Systems: Circuit connections employ dynamic color-coding to illustrate signal flow. Active connections appear in bright yellow, while inactive connections remain black, giving students an immediate and intuitive understanding of signal propagation patterns. This visual feedback reinforces conceptual learning by making cause-and-effect relationships within the circuit explicit and easy to follow.

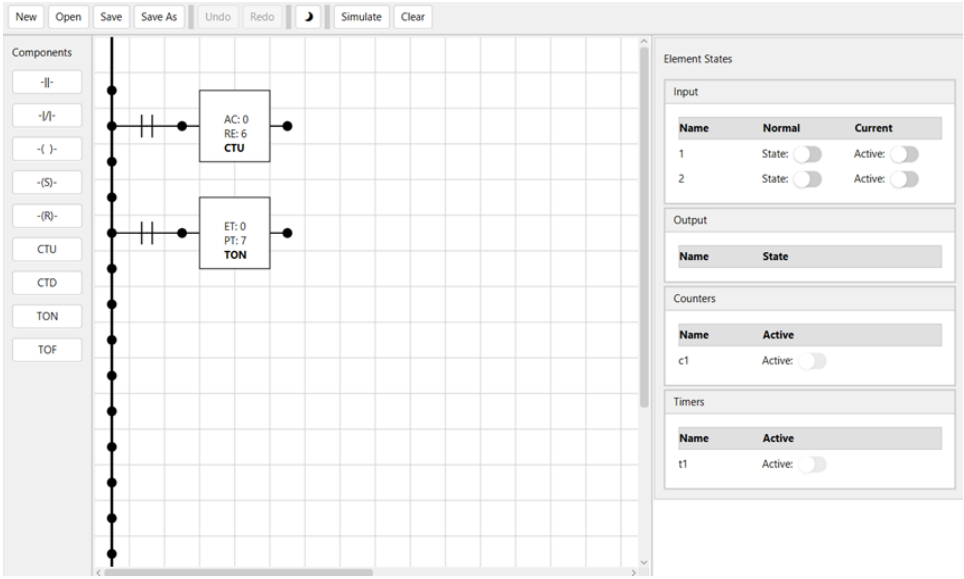


Fig. 6: User interface of the application.

4.4 Element Library and Specialized Component Implementation

LearnPLC supports a comprehensive library of standard PLC elements, each designed to balance authentic industrial behavior with educational accessibility.

Basic Elements

- **Input Contacts (NO/NC):** User-controllable circuit inputs that provide immediate visual feedback and reflect authentic logic behavior.
- **Output Coils:** Respond to input signals through name-based linking mechanisms, with clear visual indicators to reinforce understanding of output states.

Advanced Components

- **Timer Nodes (TON/TOF):** Provide ON-delay and OFF-delay functionality with configurable preset values and real-time elapsed time display.

- **Counter Nodes (CTU/CTD):** Support both count-up and count-down operations with automatic edge detection and real-time count display.
- **Set/Reset Output Nodes:** Offer latching functionality to demonstrate memory concepts in automation systems.
- **Connector Nodes:** Enable the creation of parallel circuit paths and junction points to model complex logic flows.

Educational Benefits

All components adhere to industry standards while remaining intuitive for learners. This design supports a natural progression from basic to advanced automation concepts, maintaining a consistent visual style and integrated validation mechanisms that prevent common configuration errors. As a result, students can focus on mastering core concepts without being distracted by technical complexities.

4.5 File Management and State Persistence

LearnPLC provides robust project management features, enabling students to save, share, and resume their work across multiple sessions through JSON-based serialization.

Project Data Model: The **LadderDiagramProject** class encapsulates project metadata, a complete catalog of circuit elements with configurations, a detailed connection map, and grid layout settings. The use of JSON facilitates project debugging, version control integration, and easy collaboration among students and instructors.

State Management: The **UndoRedoManager** class implements snapshot-based state management, allowing students to experiment freely without the risk of losing progress. This approach encourages exploratory learning and supports the development of practical problem-solving skills.

Error Handling: The system includes comprehensive error handling tailored to educational environments. File format validation, graceful recovery from element recreation errors, and clear, instructive error messages help students identify and correct mistakes without frustration.

Educational Impact: These implementation features demonstrate how purpose-built educational tools can offer substantial advantages over adapted industrial software. By prioritizing accessibility, ease of use, and pedagogical objectives, LearnPLC enables students to focus on learning while still experiencing an authentic PLC programming environment.

5. EDUCATIONAL VALIDATION AND EFFECTIVENESS

The platform's design integrates established educational principles and software engineering best practices to create an effective learning environment. Its implementation demonstrates how purpose-built educational tools can enhance in accessibility, usability, and learning support through thoughtful, research-informed design decisions.

5.1 Learning Progression Support and Cognitive Benefits

LearnPLC supports a natural learning progression through carefully managed complexity, aligning with key principles of educational psychology. Students begin with basic contacts and coils before progressing to timers and counters, following cognitive load theory and scaffolded learning strategies. Each new concept builds on previously mastered skills, creating a solid foundation for advanced topics.

The platform mitigates cognitive overload through multiple mechanisms:

- **Information Chunking:** Automation concepts are divided into manageable, logically connected units.
- **Visual Hierarchy:** Interface design directs attention to relevant information at each stage of learning.
- **Context-Sensitive Help:** Assistance appears only when needed, reducing interface clutter.
- **Reduced Extraneous Load:** Students focus on mastering automation concepts rather than struggling with complex software interfaces.

5.2 Real-Time Learning Reinforcement and Feedback Systems

Immediate Learning Validation:

LearnPLC delivers multiple layers of immediate feedback to enhance learning effectiveness:

- **Instant Visual Confirmation:** Color-coded connections and element states reflect the outcomes of programming decisions.
- **Logical Relationship Visualization:** Real-time updates illustrate cause-and-effect relationships within automation systems.

- **Error Detection and Correction:** Immediate identification of circuit issues supports rapid learning from mistakes.
- **Positive Reinforcement:** Clear visual confirmation of correctly functioning circuits strengthens good programming habits.

Experiential Learning Support

The platform facilitates hands-on exploration that builds practical skills:

- **Safe Experimentation:** Students can test ideas without risk to equipment or data.
- **Rapid Iteration:** Quick modification and testing encourage experimental learning.
- **Hypothesis Testing:** Students validate their understanding by building circuits and observing outcomes.
- **Problem-Solving Development:** Real-time feedback promotes systematic troubleshooting and analysis.

Metacognitive Skill Development

Monitoring and feedback mechanisms cultivate higher-order thinking skills:

- **Self-Assessment:** Students evaluate their understanding through circuit behavior.
- **Strategy Development:** Multiple solution paths encourage robust programming approaches.
- **Transfer Learning:** Skills gained with simple circuits transfer naturally to complex automation challenges.
- **Reflection:** Built-in feedback prompts students to analyze and understand circuits behavior.

5.3 Practical Application and Skill Transfer

Authentic Learning Contexts

LearnPLC immerses students in realistic automation scenarios using genuine PLC programming concepts:

- **Real-World Problem Solving:** Students design, test, and debug ladder logic circuits that reflect industrial challenges.
- **Industrial Relevance:** All functionalities and behaviors mirror those of real PLC systems.

- **Scalable Complexity:** Students progress from basic on/off control to advanced timing and counting applications.
- **Professional Practice:** The development environment parallels professional automation programming workflows.

Skill Development Outcomes

Students develop a comprehensive set of automation competencies:

- **Technical Skills:** Accurate ladder logic construction, element configuration, and circuit debugging.
- **Conceptual Understanding:** Deep comprehension of automation principles, timing relationships, and logic dependencies.
- **Problem-Solving Abilities:** Systematic approaches to troubleshooting and design.
- **Professional Readiness:** Confidence and competence transferable directly to industrial settings.

Project-Based Learning Benefits

The platform's project management system promotes extended and collaborative learning:

- **Iterative Development:** Students build complex automation solutions over multiple sessions, reflecting authentic workflows.
- **Portfolio Building:** Completed projects document learning progression for assessment and career development.
- **Collaborative Learning:** Project sharing enables peer learning and instructor feedback throughout the process.
- **Real-World Application:** Students experience the complete automation project lifecycle—from concept to testing.

6. CONCLUSIONS

This paper introduced LearnPLC, an educational platform for programmable logic controller (PLC) programming. The system employs a directed graph-based model to represent control logic, enabling efficient and accurate simulation of circuit behavior consistent with established electrical principles. The JavaFX-based user interface provides a modular and responsive environment that supports interactive learning and structured exploration of PLC concepts through visual programming.

Key contributions of this work include:

Novel Circuit Representation: A directed graph model that supports efficient simulation algorithms while preserving electrical circuit semantics essential for PLC education.

Education-Focused Design: A purpose-built interface prioritizing learning effectiveness over feature completeness, incorporating immediate visual feedback and built-in error prevention.

Modern Software Architecture: A modular and extensible design that facilitates future enhancements and customization across diverse educational contexts.

Accessibility: Free, cross-platform availability that removes common barriers to PLC education.

By balancing functionality with usability, LearnPLC demonstrates the value of purpose-built educational tools. Its emphasis on fundamental concepts and immediate feedback enables students to build a strong foundation in automation programming before transitioning to industrial platforms.

6.1 Future Research Directions

Several opportunities exist for extending LearnPLC to broaden its educational impact:

- **Enhanced Element Library:** Incorporating function blocks, mathematical operations, and communication modules would expand the range of instructional scenarios while maintaining simplicity.
- **Collaborative Learning:** Multi-user support with real-time collaboration could facilitate peer learning, instructor guidance, and remote teaching.
- **Assessment Integration:** Automated assessment and analytics could provide objective evaluation of student progress and support adaptive instruction.
- **Industrial Integration:** Introducing graduated complexity modes could bridge the gap between educational and industrial platforms, offering smoother transition pathways for learners.
- **Mobile Platforms:** Developing native mobile applications could enhance accessibility and enable ubiquitous learning opportunities.

The implications of this work extend beyond PLC education to the broader challenge of making complex technical concepts accessible to diverse learners. By applying contemporary software engineering practices, LearnPLC demonstrates how educational effectiveness can be

significantly enhanced while reducing barriers to entry. As automation technologies continue to evolve, platforms such as LearnPLC will play an increasingly important role in preparing skilled automation professionals. Its open architecture ensures continued adaptability to emerging educational needs and technological advancement.

Declarations:

Data accessibility:

There are no databases to be made public, but the codes are available upon request to the author.

Declaration of AI use:

There has been no use of AI when writing the actual paper.

Authors' contributions:

Papadhopulli I – Idea generation, study design, methodological development, data evaluation, provision of materials, review and editing, corresponding author; **Çaro M** - Idea generation, study design, experimental analysis, data evaluation, implementation, provision of materials, drafting the original manuscript.

Authors approval:

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration:

The author declares there are no conflicts of interest.

Funding:

The author declares no funding was received when writing this paper.

REFERENCES

Alves T, Buratto M, de Souza FM, Virgínia Rodrigues TV. **2014.** OpenPLC: An open source alternative to automation. In *IEEE Global Humanitarian Technology Conference*. DOI: 10.1109/GHTC.2014.6970342.

Cormen TH, Leiserson ChE, Rivest RL, Stein C, *Introduction to Algorithms*, 4th ed. Cambridge, MA: MIT Press, 2022. ISBN: 978-0262046305. DOI: 10.7551/mitpress/13329.001.0001.

Gamma E, Helm R, Johnson R, Vlissides J. 1994. *Design Patterns: Elements of Reusable Object-Oriented Software*. Reading, MA: Addison-Wesley. ISBN: 978-0201633610.

Oracle Corporation. *JavaFX Documentation*. Oracle Help Center, 2023. Available at: https://docs.oracle.com/javafx/ Last accessed 01.09.2025.

Pavković B, Živanović D. 2011. Development of a virtual environment for PLC programming and simulation. In *Proc. 34th International Convention MIPRO*. DOI: 10.1109/MIPRO.2011.5967315.

Petruzella FD. 2016. *Programmable Logic Controllers*, 5th ed. New York: McGraw-Hill Education, 2016. ISBN: 978-0073373843.

PLCLogix LLC. *PLCLogix 5000 Simulation Software*. Official product website, 2022. Available at: https://www.plclogix.com Last accessed 28.08.2025

Siemens AG. *TIA Portal: Integrated Engineering Framework*. Official product website, 2023. Available at: https://www.siemens.com/tia-portal Last accessed 20.07.2025.

Vargas H, Romero D, García F. 2023. Teaching automation with Factory I/O under a competency-based curriculum. *Multimedia Tools and Applications*, **82**, pp. 19221–19246. DOI: 10.1007/s11042-022-14047-9.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

ALBANIAN SENTIMENT ANALYSIS: EXPLORING MACHINE AND DEEP LEARNING TECHNIQUES

Alba HAVERIKU and **Elinda KAJO MEÇE**
Department of Computer Engineering, Faculty of Information Technology**,** Polytechnic University of Tirana, Albania
Corresponding author: alba.haveriku@fti.edu.al
AH, 0000-0003-2246-8408 **;** EKM, 0009-0002-7353-476
https://doi.org/10.67537/ajnts3001006

ABSTRACT

This study presents a comparative analysis of machine learning (ML) and deep learning (DL) algorithms for sentiment analysis in Albanian, a low-resource and morphologically rich language that poses distinctive challenges for natural language processing (NLP). This study explored the effectiveness of both traditional ML classifiers and neural network approaches, such as Long Short-Term Memory (LSTM) models. Experiments were conducted on two datasets: the AlbMoRe corpus and the Multilingual Twitter Sentiment Classification corpus. Given the complexity of Albanian, we examined how linguistic features influence algorithmic performance, noting that conventional ML models struggle to capture contextual sentiment effectively. We also performed a series of hyperparameter tuning experiments within the LSTM architecture, focusing on network depth, dropout rates, and embedding dimensions. These findings contribute to the advancement of NLP tools and resources for the Albanian language.

Keywords**:** sentiment analysis, machine learning, LSTM architecture, Albanian language.

1. INTRODUCTION

Albanian, a distinct branch of the Indo-European language family spoken by approximately 8 million people, is classified as a low-resource language in NLP (Hyllested and Joseph, 2022). Despite several isolated efforts to develop corpora and computational tools, most available resources remain limited in scope, proprietary, or lack comprehensive linguistic coverage. This scarcity hinders both research progress and practical NLP applications in this field.

Sentiment analysis, a prominent subfield of NLP, has gained considerable attention owing to its wide-ranging applications, including

marketing, social media monitoring, and customer experience analysis. While substantial research exists for high-resource languages such as English, French, and German, sentiment analysis for low-resource languages such as Albanian remains underexplored. The primary challenges are the limited annotated corpora, both in size and diversity, and the linguistic differences that distinguish Albanian from more widely studied languages.

This study evaluated the performance of various ML and DL algorithms for sentiment analysis tasks in Albanian. Two datasets were used in this study. First, the AlbMoRe corpus developed by Çano (2023) contains 800 manually annotated movie reviews labelled as positive or negative. Earlier experiments with this dataset (Çano 2023) applied traditional ML methods—Support Vector Machines (SVM), Logistic Regression, Decision Trees, and Random Forests—highlighting the need for further investigation using more advanced techniques.

The second dataset, compiled by Mozetič *et al*. (2016), includes over 1.6 million sentiment-annotated tweets in 13 languages, of which approximately 53,000 are Albanian. Using these two resources, we assessed the accuracy of both traditional and neural models. We also compared these results with performance on the English-language IMDb dataset (Maas *et al*. 2011), offering a cross-linguistic perspective.

The main contributions of this study are as follows:

1. A comparative evaluation of sentiment classification on both formal (movie reviews) and informal (tweets) Albanian corpora.
2. A cross-linguistic analysis comparing Albanian sentiment classification with English data, underscoring the challenges specific to low-resource languages.
3. A systematic comparison of traditional ML algorithms and DL models on low-resource data.

2. LITERATURE REVIEW: SENTIMENT ANALYSIS FOR ALBANIAN

Two of the most comprehensive corpora developed for the Albanian language are the Albanian National Corpus, created by researchers at Saint Petersburg State University (Arkhangelskiy, 2012), and the sqGlobe Corpus, developed at Beijing Foreign Studies University (Jing *et al*. 2012). The Albanian National Corpus contains approximately 16.6 million tokens, whereas sqGlobe comprises approximately 1 million words

annotated with part-of-speech tags and lemmatization. Despite their size and richness, both corpora are closed-source and thus unavailable for broader linguistic research purposes.

Within the Universal Dependencies (UD) initiative (Nivre *et al*. 2017), three treebanks currently exist for Albanian. The Treebank of Standard Albanian (TSA) consists of 60 sentences and 922 tokens annotated with part-of-speech and syntactic information (Toska *et al*. 2020). The Saarbrücken Treebank of Albanian Fiction (STAF) includes 202 sentences with 3,499 tokens (Talamo, 2025). The third, Gheg Pear Stories (GPS), was specifically built for the Gheg dialect and contains 966 sentences comprising 15,990 tokens (Ebert, 2022).

Kote *et al*. (2019) introduced a morphologically annotated corpus containing approximately 118,000 tokens. In a more recent contribution, Kote *et al*. (2024) presented the SALT (Standard Albanian Language Treebank), which consists of 1,400 sentences and 24,537 tokens. This treebank provides comprehensive annotations, including sentence and word segmentation, as well as Universal Part-of-Speech (UPOS) tags.

Having established a general overview of the major linguistic corpora available for Albanian, we now turn to the research on sentiment analysis and opinion mining. One of the earliest studies on sentiment analysis for Albanian was conducted by Biba and Mane (2014), who developed and evaluated a set of machine learning classifiers (SVM, Logistic Regression, Naïve Bayes) on a corpus of 400 political news articles. Their experiments reported accuracy rates between 86% and 92%, depending on the topics. Subsequently, Kote *et al*. (2018) presented an experimental setup comparing machine learning algorithms for opinion mining, achieving initial results ranging from 79% to 94% accuracy. In a later study, they trained and tested 50 different classification algorithms in Weka by combining data from five different topics.

Skënduli *et al.* (2018) implemented CNN-based models for emotion detection in Albanian microblogging texts, achieving accuracies above 90% on some subset corpora. Besjana *et al*. (2024) analyzed the sentiment of news articles and its role in predicting fake news. By including sentiment analysis as an attribute in their model, the authors improved the performance of the XGBoost model for fake news detection by 0.4%.

3. METHODOLOGY

This section outlines the methodological framework used to classify Albanian texts according to their expressed sentiments, categorizing each

instance as either positive or negative. The classification process comprised four main phases: i) data collection, ii) data preprocessing, iii) model building, and iv) model evaluation.

3.1. *Data collection & Pre-processing*

The first step in our pipeline involved collecting the data necessary for this study. We employed three corpora: two containing Albanian texts with sentiment annotations and one English-language corpus for comparative analysis. Table 1 presents the details of each dataset, encompassing both formal and informal language sources.

Table 1. Datasets used for model training and evaluation

Name of Corpus	Number of records	Language	Sentiment
AlbMoRe (Çano *et al.* (2023))	800 movie reviews	Albanian	positive, negative
Mozetic et al. (2016)	53,000 tweets	Albanian	positive, negative, neutral
IMDb (Maas *et al.* (2011))	25,000 movie reviews	English	positive, negative

The preprocessing phase is a critical step that ensures that raw text data are standardized and cleaned, making them suitable for feature extraction and model training. Our preprocessing pipeline involved several stages designed to remove noise and inconsistencies from the data set. Specifically, we applied the following procedures:

- **Stopword Removal:** Albanian stopwords were removed using a custom-built list to ensure that frequent but uninformative words did not dominate the feature space.
- **Noise Removal:** URLs and hashtags were removed from the datasets because they carried little semantic value relevant to sentiment.
- **Text Normalization:** All text was converted to lowercase to reduce dimensionality and improve consistency.
- **Label Filtering:** In the Mozetič *et al*. (2016) dataset, which originally contained three sentiment labels (positive, negative, and neutral), we excluded all records that were labeled as neutral. This allowed us to focus on a binary classification task that distinguished only between positive and negative sentiments.

- **Punctuation Removal:** All punctuation marks were removed from the datasets to simplify the feature space. Although punctuation can carry affective or stylistic cues relevant to sentiment, we opted for its removal in this study and suggested that future research explore punctuation as a potential feature for sentiment classification in Albanian.

This pipeline reflects a balance between linguistic sensitivity and computational efficiency, making it especially adaptable to low-resource languages, such as Albanian.

3.2. Model Building and Evaluation

In the model development phase, we evaluated both traditional machine learning algorithms and deep learning approaches using Long Short-Term Memory (LSTM) networks. For feature extraction, we applied the Term Frequency–Inverse Document Frequency (TF-IDF) technique, which quantifies the importance of words within a document relative to the corpus. TF-IDF was chosen for its computational efficiency, interpretability, and widespread use in sentiment analysis.

The processed datasets were split into training and testing subsets using an 80–20 ratio, with 80% of the records allocated for training and 20% for testing. We used the *random state* hyperparameter to control the randomness and ensure the reproducibility of the train–test splits across multiple executions.

Traditional machine learning models were implemented using the *Scikit-learn* library. For the LSTM architecture, we employed the *TensorFlow* and *Keras* frameworks. Initially, the LSTM model was tested using the default hyperparameters across all datasets. We then fine-tuned the LSTM model specifically for the AlbMoRe dataset, which yielded more consistent and accurate results. The tuned parameters included the: dropout rate in the LSTM layer, number of epochs, and batch size during model training. The impact of each hyperparameter adjustment on the model performance is presented in Table 1.

To assess the model performance, we used the *predict()* function from scikit-learn to generate predictions for the test data. The evaluation was based on two widely adopted metrics for classification tasks, accuracy and precision, which were selected to provide a balanced view of the models' performance. The results for each model are listed in Table 2.

4. RESULTS AND DISCUSSIONS

Table 2 presents the evaluation results for each algorithm across the different corpora. In this table, the terms “Corpus 1,” “Corpus 2,” and “Corpus 1+2” refer to the AlbMoRe, Mozetič *et al*. (2016) corpus, and their combined version, respectively.

Overall, the models performed best on the AlbMoRe corpus, achieving the highest accuracy and precision values among all the algorithms. In particular, the Support Vector Machine (SVM) algorithm obtained the highest accuracy of 94.4%. In contrast, models trained and tested on the Mozetič *et al*., (2016) corpus exhibited lower accuracy than those using the AlbMoRe and IMDb corpora. This reduction in performance may be attributed to less reliable sentiment annotations in the Albanian subset of the Mozetič corpus, which likely introduced noise and inconsistencies during training and evaluation. Additionally, the informal and short-text nature of social media content (e.g., Twitter) complicates accurate sentiment detection.

Notably, the performance of the models trained on Corpus 1+2 closely resembled that of Corpus 2 alone. This result can be explained by the large difference in size between the two corpora, whereby the substantially larger Mozetič corpus effectively dominates the combined dataset, thereby diminishing the influence of the smaller AlbMoRe corpus during training.

Table 2. Accuracy and precision metrics across algorithms and corpora

	SVM		Naïve Bayes		Logistic Regression		LSTM	
	Accuracy	Precision	Accuracy	Precision	Accuracy	Precision	Accuracy	Precision
Corpus 1	94.4%	0.94/0.95	93.1%	0.90/0.96	90.0%	0.89/0.91	91.3%	0.90/0.92
Corpus 2	79.6%	0.81/0.6	78.3%	0.78/0.73	78.5%	0.79/0.6	75.9%	0.84/0.48
Corpus 1+2	79.5%	0.81/0.66	77.9%	0.78/0.73	78.4%	0.79/0.73	76.0%	0.83/0.50
IMDb	89.4%	0.89/0.90	86.3%	0.87/0.85	89.1%	0.88/0.90	88.1%	0.87/0.89

Furthermore, we explored the effects of several hyperparameter settings on the LSTM model trained on the AlbMoRe corpus. The impact of each hyperparameter is reported in Tables 3-5, where only one parameter was varied at a time, while all others were held at their default values.

As shown in Table 3, increasing the batch size gradually improved the accuracy, with the highest accuracy (91.3%) achieved at a batch size of 128. Similarly, increasing the number of epochs (Table 4) produced consistent gains in performance, with the highest accuracy of 91.88% being reached after 10 epochs.

Table 5 shows the effect of dropout rate on LSTM accuracy. Dropout values of 0.1 and 0.4 yielded the highest observed accuracies (91.87% and 91.8%, respectively), whereas intermediate dropout values produced slightly lower performances. These results suggest that both smaller and larger dropout rates can help regularize the model without excessively diminishing its capacity to learn the patterns.

Overall, these findings demonstrate that both the algorithm choice and dataset quality exert a strong influence on sentiment classification performance. In particular, the AlbMoRe corpus supports more robust training outcomes, and careful hyperparameter tuning of the LSTM model can provide additional performance gains in deep learning–based sentiment classification tasks.

Table 3. Effect of Batch Size on LSTM Accuracy

Batch size	**16**	**32**	**64**	**128**
Accuracy (%)	**90.0**	**90.06**	**90.6**	**91.3**

Table 4. Effect of Number of Epochs on LSTM Accuracy

Number of Epochs	**5**	**6**	**8**	**10**
Accuracy (%)	**90.6**	**91.3**	**91.87**	**91.88**

Table 5. Effect of Dropout Rate on LSTM Accuracy

Dropout Rate	**0.1**	**0.2**	**0.3**	**0.4**
Accuracy (%)	**91.87**	**90.6**	**90.6**	**91.8**

5. CONCLUSIONS

This study presents a comparative analysis of several widely used machine learning algorithms, such as SVM, Naïve Bayes, and Logistic Regression, alongside an LSTM neural architecture for sentiment analysis in the Albanian language. Experiments were conducted across multiple corpora, including AlbMoRe, the Mozetič *et al*. (2016) dataset, and the English IMDb dataset, to evaluating model performance in terms of accuracy and precision.

The findings show that the AlbMoRe corpus consistently yielded the best performance, outperforming all other datasets. This result is largely attributed to the higher quality of annotation and well-structured sentences in AlbMoRe, as opposed to the less consistently labelled content found in the Mozetič dataset.

In addition to the baseline evaluations, hyperparameter tuning experiments for the LSTM model using the AlbMoRe dataset produced promising results. Optimized configurations of these parameters led to noticeable improvements in the classification performance. These findings highlight the potential of deep learning methods for Albanian sentiment analysis, emphasizing the importance of model optimization and dataset quality.

Future work may involve extensive hyperparameter searches and experimentation with more advanced deep learning architectures, such as Bidirectional LSTM, GRU, or transformer-based models such as BERT. Another important avenue is to increase the size and diversity of Albanian sentiment-related corpora by incorporating different dialects and domains (e.g., social media, literature, and news).

In conclusion, this study contributes to the growing body of research on Albanian as a low-resource language and provides evidence that, despite linguistic complexity and limited resources, acceptable performance can be achieved through careful corpus selection, algorithm testing and model fine-tuning.

DECLARATIONS

Declaration of AI use: We confirm that no AI-assisted technologies were used to create this article.

Originality and Authorship: This work is based on the original data and ideas developed by the authors.

Author contributions

Haveriku A: Ideation, conceptualization, methodology, data interpretation, resources, writing—original draft preparation, and corresponding author; **Kajo Meçe** E: Ideation, conceptualization, methodology, data interpretation, resources, review and editing, supervision.

Authors approval:

All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare there are no conflicts of interest.

Funding: There is no related funding for this paper.

REFERENCES:

Arkhangelskiy T, Michael D, Yurievich RA, Alexandrovna MM. 2012. *Korpusi i gjuhës shqipe: Drejtimet kryesore të punës*. https://api.semanticscholar.org/CorpusID:185412758.

Biba M, Mane M. 2014. Sentiment analysis through machine learning: An experimental evaluation for Albanian. In: *Recent advances in intelligent informatics* (pp. 195–203). Springer.

Çano E. 2023. *AlbMoRe: Movie reviews in Albanian* [Corpus]. LINDAT/CLARIAH-CZ digital library, Charles University. http://hdl.handle.net/11234/1-5165.

Ebert C, Islamaj A, Kuqi A, Sonnenhauser B, Plamada M, Widmer P. 2022. UD Gheg Pear Stories [corpus].

Hyllested A, Joseph BD. 2022. Albanian. In J. Clackson (Ed.), The Cambridge handbook of Indo-European Languages (pp. 223–245). Cambridge University Press.

Jing K, Qiao J, Shihao Y, Tong H, Wang YF, Wang X, Hu Z, Lacmi E, Allmetaj E, Chen T, Zhang W, Zhang H, Lu Y, Ai W. 2012. The SQGlobe corpus: A balanced 1M-word contemporary written Albanian corpus, lemmatized and POS-tagged [corpus].

Kote N, Biba M, Kanerva J, Rönnqvist S, Ginter F. 2019. Morphological tagging and lemmatization of Albanian: A manually annotated corpus and neural models. arXiv preprint arXiv:1912.00991. http://arxiv.org/abs/1912.00991.

Skënduli MP, Biba M, Loglisci C, Ceci M, Malerba D. 2018. User-emotion detection through sentence-based classification using deep learning: A case-study with microblogs in Albanian. In: *International Symposium on Methodologies for Intelligent Systems* (pp. 258–267). Springer.

Kote N, Biba, M, Trandafili E. 2018. A thorough experimental evaluation of algorithms for opinion mining in Albanian. In International Conference on Emerging Internetworking, Data & Web Technologies (pp. 525–536). Springer.

Kote N, Rushiti R, Çepani A, Haveriku A, Trandafili E, Meçe Kajo E, Skënderi Rakipllari E, Xhanari L, Deda A. 2024. Universal Dependencies treebank for Standard Albanian: A new approach. In Proceedings of the Sixth International Conference on Computational Linguistics in Bulgaria (CLiB 2024) (pp. 80–89). Institute for Bulgarian Language, Bulgarian Academy of Sciences. https://aclanthology.org/

Maas AL, Daly RE, Pham PT, Huang D, Ng AY, Potts C. 2011. Learning word vectors for sentiment analysis. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologi*es (pp. 142–150). Association for Computational Linguistics. https://aclanthology.org/P11-1015.

Mozetič I, Grčar M, Smailović J. 2016. *Twitter sentiment for 15 European languages* [corpus]. Slovenian language resource repository CLARIN.SI. http://hdl.handle.net/11356/1054.

Muraku B, Xiao L, Meçe EK. 2024. Toward detection of fake news using sentiment analysis for Albanian news articles. In L. Barolli (Ed.), Advances in Internet, Data & Web Technologies (EIDWT 2024), Lecture Notes on Data Engineering and Communications Technologies (Vol. 193, pp. 665–676). Springer. https://doi.org/10.1007/978-3-031-53555-0_55**Nivre J, Zeman D, Ginter F, Tyers F. 2017**. Universal Dependencies. In A. Klementiev & L. Specia (Eds.), *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts*. Association for Computational Linguistics. https://aclanthology.org/E17-5001/.

Talamo L. 2025. Introducing STAF: The Saarbrücken Treebank of Albanian Fiction. Zenodo. https://doi.org/10.5281/zenodo.14552809.

Toska M, Nivre J, Zeman D. 2020. Universal dependencies for Albanian. In Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020). Association for Computational Linguistics.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

PHYTOPLANKTON MONITORING AND GENETIC ANALYSIS OF LAKE BUTRINTI: ADVANCED METHODS AND ENVIRONMENTAL IMPLICATIONS

Xhensila OMERI

Department of Biotechnology, Faculty of Natural Sciences, University of Tirana, Albania

Ariola BACU

Department of Biotechnology, Faculty of Natural Sciences, University of Tirana, Albania

NanoAlb Center, Albanian Academy of Sciences

Author for correspondence: xheniomeri@gmail.com

XhO, 0009-0005-3937-6245 ; AB, 0000-0001-6390-2632

https://doi.org/10.67537/ajnts3001007

ABSTRACT

Lake Butrinti, located near Saranda in southern Albania, is a unique brackish lagoon shaped by tectonic activity and characterized by significant ecological, historical, and economic value. Its distinct nature has sparked scientific interest in understanding the factors that influence water quality and the lagoon's capacity to support diverse life forms. This paper reviews the findings on water quality based on physical (temperature, turbidity, salinity, and pH), chemical (oxygen, OD, COD, BOD, nitrites, nitrates, phosphates, heavy metals, pesticides, and organic matter), and biological indicators (phytoplankton chlorophyll a, Carlson Trophic State Index-TSIC, biomass, and cytotoxicity testing via bio-reporting bacteria). The data reported data were organized to depict annual, seasonal, and spatial dynamics more clearly. Furthermore, the methodologies applied to assess phytoplankton diversity, including microscopy, chemotaxonomy, specific-PCR, CARD-FISH, and sequencing, are reviewed, with a discussion of their respective advantages and limitations. The findings indicate that anthropogenic activities, including past interventions in the hydrological network, fishing practices, commercial aquaculture, and pollution, along with the effects of climate change, have negatively impacted the lagoons on phytoplankton diversity, and their advantages versus limitations were discussed. In conclusion, anthropogenic activities (former interventions in the hydrological network, fishing, commercial aquaculture farming, and water pollution) have negatively impacted the lagoon's physical, chemical, and biological integrity. To prevent significant ecological and

socio-economic consequences, such as dystrophic crises, reduced fishing and mussel production, ecological degradation, biodiversity loss, and decreased aesthetic value, there is a pressing need for more comprehensive and systematic research. Such research should employ both conventional and advanced methods to monitor changes, mitigate adverse effects, and safeguard the ecological integrity and economic sustainability of the lake.

Keywords: Albanian lagoons, physic-chemical parameters, algal bloom, analytical methods

1. INTRODUCTION

Lake Butrinti is a unique brackish lagoon shaped by tectonic activity and is characterized by significant ecological, historical, and economic importance. Covering an area of approximately 16.3 km² with an average depth of 14 m (Topi *et al*. 2013c; Guri *et al*. 2014), it receives freshwater from the Pavllo and Bistrica rivers and saltwater from the Ionian Sea through the Vivari Canal. The lake is also connected to Lake Bufi via the Rreza/Bufi Canal, with the surrounding bedrock primarily composed of limestone (Hounslow and Chepstow-Lusty, 2004). Subterranean springs emerge along the eastern lakeshores, and two pumping stations discharge drainage waters from the Vurgu and Vrina plains at Manastiri (northern shore) and Dajlani (southern shore), respectively (Miho *et al*. 2013). The lagoon is a vital ecosystem supporting rare habitats and biodiversity (Miho, 1994; Miho and Witkowski, 2005; Bego and Malltezi, 2010; Bushati, 2013; Miho *et al*. 2013), with the cultivation of *Mytilus galloprovincialis* playing a key role in the local economy (Bego and Malltezi, 2010; Miho *et al*. 2013). The area holds outstanding cultural value and is protected under national and international legislation, including its designation as a UNESCO World Heritage Site (UNESCO, Topi *et al*. 2013b). Moreover, Lake Butrinti forms a core component of the Ramsar transitory-marine area of Çuka Channel-Butrinti-Stillo Cape (13,500 ha; DCM 531/2002; Site No.1290, 28.3.2003).

However, increasing anthropogenic pressures, including the diversion of the Bistrica River in the late 1950s, waste disposal, and growing tourism, have exacerbated environmental stress (Miho 1994; Xhulaj *et al*., 2008). River diversion and reclamation of the Vurgu Plain significantly reduced the surface area of Lake Butrinti and its associated wetlands, decreased freshwater input, and led to increased mineralization of the lagoon. Consequently, Butrinti is classified as a meromictic lake with long-

standing water stratification (epilimnion and hypolimnion) and a history of anoxic conditions that facilitate hydrogen sulfide (H_2S) accumulation, limiting benthic life (Pano, 1984; Miho, 1994; Moisiu *et al*. 2016; Bacu and Zaho, 2022).

2. MATERIALS AND METHODS

This review synthesizes findings on water quality based on physical (temperature, turbidity, salinity, and pH), chemical (oxygen, OD, COD, BOD, nitrites, nitrates, phosphates, heavy metals, pesticides, and organic matter), and biological parameters (phytoplankton chlorophyll a, Carlson Trophic State Index – TSIC, biomass, and cytotoxicity testing using bio-reporting bacteria). Data were collected from studies conducted over the past decades, including Pano (1984), Miho (1994), Hounslow and Chepstow-Lusty (2004), Miho and Witkowski (2005), Xhulaj *et al*. (2008), Bego and Malltezi (2010), Miho *et al*. (2013), Bushati (2013), Topi *et al*. (2013 a), Guri *et al*. (2014), Moisiu *et al*. (2016), and Bacu and Zaho (2022), among others. The reviewed data are presented in a structured format to facilitate a clearer depiction of the annual, seasonal, and spatial dynamics of these parameters. Additionally, methodologies previously employed for phytoplankton diversity assessment, including microscopy, chemotaxonomy, specific PCR, CARD-FISH, and sequencing, are examined. Their respective capacities and limitations are critically discussed, drawing on findings from Bacu *et al*. (2022), Bacu and Zaho (2022), Bacu *et al*. (2024), Omeri *et al*. (2024), and other relevant studies.

An overview on physical and chemical characteristics of waters and sediments

Physical characteristics

Butrinti region has a typical Mediterranean climate, which is reflected in both water temperature (Table 1) and precipitation patterns (Dedej and Bino, 2003; Topi *et al*. 2013(c); Velaj, 2015).

Table 1. Epilimnion temperature (˚C) of Lake Butrinti

References	Year of measurement	Winter-autumn	Spring-summer
Osmani Miri and Peja, 2012	2012	9˚C	28˚C

Bushati, 2013	2006-2010	6˚C	28˚C
NEA 2015	2013	19.3˚C	24.2˚C

Rainfall is abundant, with higher levels occurring in autumn and winter and lower levels in spring and summer. The annual average precipitation is approximately 1,500 mm (Miho, 1994; Dedej and Bino, 2003; Bego and Malltezi, 2010; Zotaj, 2010; Topi e*t al*. 2013a; Pano, 2015; Velaj, 2015).

The tidal cycle at Lake Butrinti follows a six-hour periodicity (Pano, 2015). However, hydrometeorological factors, including components of the water balance, wind regime, and wave dynamics, also influence water level oscillations. These factors do not affect not the pattern of fluctuations, but their amplitude.

Owing to its connection with the Ionian Sea, Lake Butrinti is influenced by marine winds, with average speeds of 3.2–3.5 m/s in winter and 2.2–2.3 m/s in summer (Bego and Malltezi, 2010; Topi *et al*. 2013c; Zotaj, 2010).

Water transparency (Table 2) varies in response to rainfall, wind intensity, and plankton growth (Miho, 1994; Bushati, 2013; Miho *et al*. 2013). The highest transparency values are generally recorded during winter, late spring, and early summer (Bushati, 2013).

Table 2. Transparency of the Lake Butrinti based on Secchi disc measurements (m)

References	Depth (m)
Miho, 1994; Bushati, 2013	0.8/1 - 4
Pano, 2015	0.8 - 3.2
Heywood, 2017	2.8 - 3
Çako *et al.* 2013	1.85 - 4.32

The sediment archive of Lake Butrinti provides evidence of variations in the Mediterranean climate and anthropogenic impacts on the lake's trophic status and hydrology. Lake sediments are excellent recorders of environmental change, capturing signals from within the basin and from the surrounding catchment. When coupled with robust chronological constraints, lacustrine sediment archives can effectively reveal climatic, human, and tectonic influences on the environment (Xhulaj *et al*. 2008; Ariztegui *et al*. 2010). Sediments consist primarily of carbonates, organic

matter, and clay layers (Morellón *et al*. 2016) and are characterized by finely stratified silt resulting from the seasonal deposition of calcite, which is rich in organic materials and clays. The presence of heavy metals in lake sediments has also been investigated by Topi *et al*. (2012) and Omeri *et al*. (2024).

Organic carbon (Corg) levels in the sediments exhibited an average value of 1.792 ± 0.336% dry weight (d.w.), while total nitrogen (TN) averaged 0.197 ± 0.043% d.w. Their spatial distribution is relatively uniform between the northwestern and southeastern regions of the lagoon (Moisiu *et al.* 2016).

Chemical characteristics

The salinity of Lake Butrinti is influenced by its connection with the Ionian Sea through the Vivari Canal, contact with Lake Bufi, freshwater inflows from the Bistrica and Pavllo rivers, and precipitation and evaporation processes (Miho *et al*. 2013; Pano, 1984; 2015). Reported salinity values vary among authors (Tables 3 and 4); however, a consistent finding is that salinity increases with depth (Pano, 1984; Miho, 1994; Pano, 2015; Moisiu *et al*. 2016). Seasonal variation is also evident, with higher salinity values observed during summer compared than in winter (Pano, 1984; Dedej and Bino, 2003; Miho *et al*. 2013). In the epilimnion, salinity ranges from 14–35‰, reaching maximum levels in late summer and minimum levels in winter. Below a depth of 6 m, salinity remains consistently higher, between 20–35‰. In contrast, salinity in the adjacent seawaters is significantly higher, ranging from 30.4‰ at the surface to 40.9‰ at a depth of 5 m (Miho, 1994; Bushati, 2013; Miho *et al*. 2013).

Table 3. Salinity (‰) in hypolimnion and epilimnion

References	Hypolimnion (under 6 meter)	Epilimnion (surface waters)
Miho, 1994	35.1‰	18‰
Anonymous, 2010; Miho *et al.* 2013	20-35‰	14-35‰
Moisiu *et al.* 2016	35‰	24.3‰

Table 4. Salinity (‰) in summer and winter

References	Summer	Winter
Pano, 1984	26‰	13‰
Dedej and Bino, 2003	33‰	15‰

pH

The waters of Lake Butrinti are slightly alkaline (Table 5). Generally, alkalinity is lower during summer compared to winter, as reported by Dedej and Bino (2003), Miho *et al.* (2013), and Heywood (2017), with some exceptions noted by Pano (1984). Regarding vertical distribution, pH values tend to decrease from the surface to greater depths (Pano, 1984; Miho, 1994; Bushati, 2013).

Table 5. Epilimnion pH values during winter and summer at Lake Butrinti

References	Summer	Winter
Pano 1984; Miho 1994; Bushati 2013	8.7	7.9
Miho *et al.*, 2013; Dedej and Bino 2003	6.5	9.5
NEA, 2015	7.18	8.53
Çako *et al.*, 2013	7.8	8.08

Oxygen, COD, BOD and TOC

The oxygen levels in Lake Butrinti are influenced by multiple factors, including gas exchange, photosynthesis (Pano, 2015), wind, and temperature (Pano, 1984; Gabeira, 2023). Seasonal trends show that dissolved oxygen (DO) concentrations are generally higher in winter than in summer (Pano, 1984; Miho, 1994; Miho *et al*. 2013), and that DO levels decrease with depth (Table 6). Miho *et al*. (2013) stated that epilimnion waters are consistently well oxygenated, often exhibiting saturation levels above 100%, with maxima recorded in late winter and early spring. However, below depths of 4–5 m, oxygen levels decline to less than 50%,

and at 6–8 m, they drop further, oscillating between 0 and 2 mg/L. In contrast, the hypolimnion is characterized by the presence of hydrogen sulfide (H_2S) and the absence of oxygen (Miho, 1994; Dedej and Bino, 2003; Bushati, 2013; Pano, 2015; Velaj, 2015).

Table 6. Oxygen level in different depths of Lake Butrinti

References	Epilimnion	5 m	6-8 m	10 m and below
Miho *et al.* 2013; Bushati 2013	5 mg/l	14 mg/l	0-3 mg/l	0
Dedej and Bino, 2003	8-9 mg/l		0	0

The chemical oxygen demand (COD) in Lake Butrinti ranged from 1.52 to 9.04 mg/L across the surface, mid-depth, and bottom layers, indicating poor water quality according to the Norwegian Water Institute (NIVA) classification (6.5–15 mg/L) (Heywood, 2017). The biological oxygen demand (BOD) ranged from 0.5 to 1.25 mg/L, reflecting moderate water quality. For 2023–2024, the maximum COD values were recorded in September, with minimum values observed in April. BOD peaked in November and reached its lowest value in March (Omeri *et al*. 2024). Organic matter concentrations ranged from 2 to 17 mg/L (Dedej and Bino, 2003), with higher levels in April than in August, suggesting that phytoplankton are the principal contributors to organic matter production during the spring season (Miho *et al*. 2013).

Nutrients (N and P)

Phosphorus is a key limiting nutrient for aquatic organisms; at elevated concentrations, it can be harmful by promoting eutrophication and algal blooms in coastal transitional waters such as Lake Butrinti. Consequently, phosphate levels in natural waters are among the most critical environmental concerns. Many regulations focus on limiting total phosphorus concentrations to prevent harmful algal blooms and maintain water quality, with typical standards ranging from 0.01 - 0.10 mg/L (Boyd 2001). However, data on nutrient content (N and P) in Lake Butrinti are limited and sporadic. Heywood (2017), based on the Official Environment Report of Albania (2014), reports total phosphorus concentrations from 0.03 to 0.35 mg/L during monitoring in 2013 (two seasonal phases). Nitrate ($N–NO_3$) values ranged from 0.1 to 1.6 mg/L, whereas nitrite (N–

NO_2) values ranged from 0.01 to 0.02 mg/L. Despite limited data, the phosphate concentrations often exceed 0.1 mg/L.

Kolitari *et al*. (2013) documented a significant increase in nitrite and phosphate levels between 2004 and 2011, with nitrite peaking in June and phosphate in October. High nutrient concentrations, combined with elevated temperatures, evaporation, limited water exchange, and low flow conditions, particularly in summer, likely contribute to dystrophic crises and periodic mussel die-offs. This situation may worsen due to nutrient leaching from increasing agricultural and urban activities in the catchment areas, including the Vurgu and Vrina plains and the urban centers of Ksamil and Manastiri, where untreated waters are discharged into the lake via the Manastiri and Dajlani pumping stations.

Table 7. Nutrient (N and P) presence at surface waters of Butrinti

	Heywood 2017	Kolitari *et al.*, 2013	Pano 2015
Nitrites	Low	High	Low
Nitrates	Low	-	0
Phosphates	High	High	-

Sulfides and hydrogen sulfide

Miho *et al*. (2013) reported that sulfides and hydrogen sulfide (H_2S) are consistently present in Lake Butrinti waters below 7 m depth. In some cases, particularly during hot summers, sulfides are also detected in surface waters, causing crises among sedentary biota, especially mussels. Sulfate concentrations were relatively high, reaching up to 2.86 mg/L, with a vertical distribution similar to that of salinity, and the lowest concentrations were observed in the surface layers (up to 2.01 mg/L). The presence of H_2S is attributed to sulfur-reducing bacteria that can reduce acidic sulfur salts under anaerobic conditions. Its concentration increases with depth, reaching values exceeding 5.0 mg/L at the lake bottom (Miho, 1994; Dedej and Bino, 2003; Pano, 2015; Velaj, 2015).

Heavy metals

Few studies have addressed the presence of heavy metals in the waters and sediments of Lake Butrinti (Topi *et al*. 2012; 2013a,b). Lead (Pb) concentrations were higher in both the hypolimnion and epilimnion during winter, while cadmium (Cd) is detected at both the surface and depth

during autumn. Chromium (Cr) predominantly occurred at depth, with peak values observed in the summer. Copper (Cu) is present throughout the water column, with higher concentrations in the summer. Mercury (Hg) has also been detected. Overall, the annual heavy metal concentrations generally fell within the normal range, with the notable exception of mercury. The average Hg concentration across all seasons (0.79 µg/L) exceeded the EU guideline of 0.05 µg/L (Bushati, 2013; Table 8). More recent studies (Omeri *et al*. 2024) have reported that Cd, Cr, and Pb levels during 2023–2024 also surpassed the EU reference values.

Table 8. Heavy metal concentrations in Lake Butrinti waters compared to EU standards (µg L^{-1}) (Topi *et al*. 2012; 2013a,b; Omeri *et al*. 2024)

Reference	Cd	Cr	Cu	Pb	Hg
Topi *et al.* 2013	0.078	1.41	8.26	1.5	0.79
Topi *et al.* 2012	0.075	11.06	20.15	1.78	0.23
Omeri *et al.* 2024	1.4	41	18	29	0.001
EU Standard	**1**	**20**	**50**	**7.2**	**0.05**

Pesticides

Nuro and Marku (2011) reported the presence of pesticides in the waters of Butrinti, including DDT, Lindane, HCB, Aldrins, and Heptachlors. The concentration of organochlorine pesticides ranged from 7.3 to 30.7 ng/L, whereas the general EU limit for pesticide residues (EQS 2013/39/EU) is approximately 0.01 ng/L (0.00001 µg/L). These values indicate that pesticide levels in Butrinti waters substantially exceed the EU standards. This finding provides further evidence of the adverse impact of agricultural and urban activities in the surrounding areas. Given that such activities have intensified in recent years; it is reasonable to suspect that the associated environmental impact may now be even greater.

Investigation of the toxicity of waters at Butrinti Lake using bio-reporting bacteria

Physicochemical analyses of waters and sediments of Butrinti Lake confirmed the presence of contaminants that may affect the organisms inhabiting the ecosystem (Fig. 1a) and the adjacent marine protected area. While the measurement of individual chemical parameters provides valuable insight into water quality, assessing cytotoxicity using

luminescent bio-reporting bacteria allows for the simultaneous evaluation of the combined effects of all pollutants. This was achieved by monitoring the changes luminescence emitted by the bacteria (Fig. 1b). Biosensors based on luminescent bacteria are highly sensitive to toxic substances and are valuable tools for monitoring water safety and chemical quality (Bond and Martin, 2005; Woutersen *et al*. 2010).

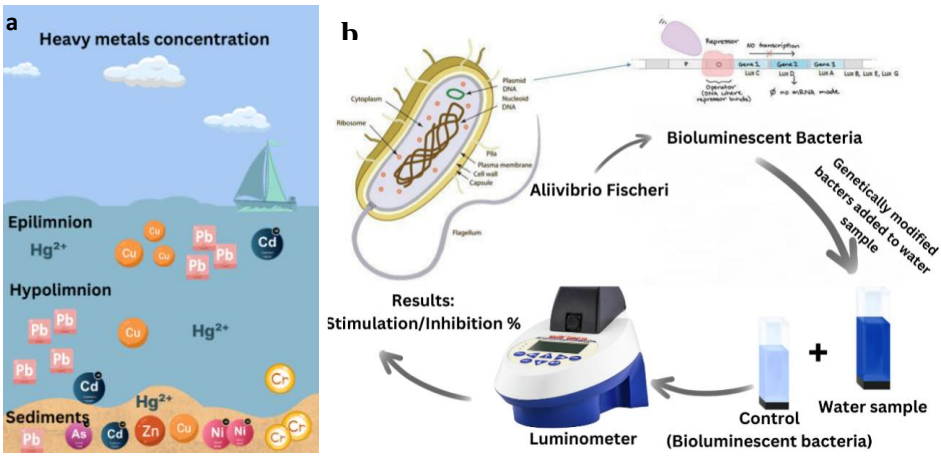


Fig. 1: (*a*) HM distribution in the epi/hypolimnion and sediments of Butrinti Lake based on previously reported data. **(*b*)** Use of bioluminescent bio-reporting bacteria for the detection of cytotoxicity. Figures were prepared using the Canva software.

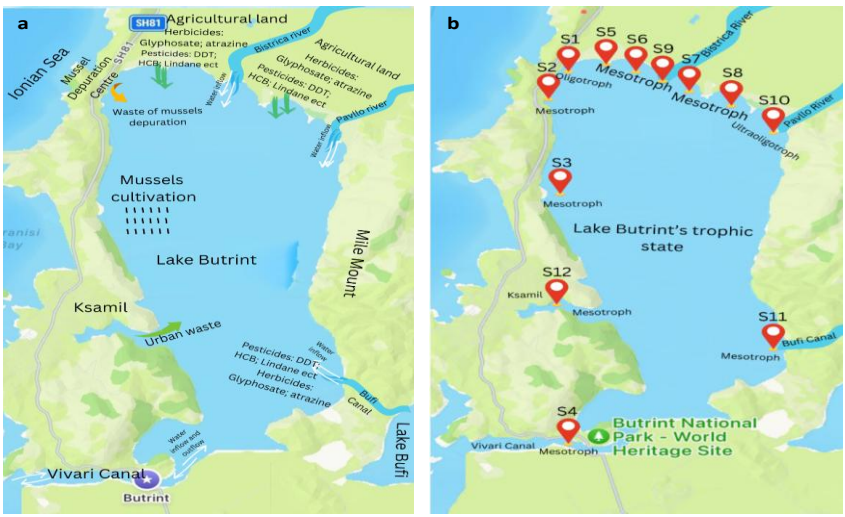
Water samples from 12 stations in Butrinti Lake were analysed for cytotoxicity (Omeri *et al*. 2024) using luminescent *Vibrio fischeri* according to ISO 11348-3:2007. The results revealed the impact of various categories of pollutants, including pesticides, heavy metals, urban waste, and aquaculture by-products, together with the hydrological characteristics of the lake, on bacterial toxicity (Fig. 2). Data on heavy metal concentrations in sediments (Bacu *et al*. 2024; Fig. 2d) and water (Table 8) were used to interpret the cytotoxicity levels at the corresponding stations (Fig. 2a–d). The findings indicate that the hydrological regime plays a significant role in determining water cytotoxicity. Although Cd, Cr, and Pb concentrations in water indicated a "Moderately Polluted" status (according to EU standards; Table 8), and Cr and Ni in sediments range from "Moderately Polluted" (Stations 6–10) to "Heavily Polluted" (Stations 4, 6–11) (according to USEPA, 1991 Guidelines), the Cytotoxicity Induction Factor (IF) was significantly high only at Station 2.

This location corresponds to the Vivari Canal, an area characterized by high fluctuations in inflow and outflow between Butrinti Lake and the Ionian Sea. Moreover, stations situated near agricultural areas exhibited IF values ranging from 1.5 to 2, highlighting the role of agricultural pollutants in the degradation of water quality.

Phytoplankton data

Butrinti Lake is considered one of the most important transitional ecosystems along the Albanian coast, offering substantial ecological and economic benefits, as well as high biodiversity and productivity. The lagoon is extensively used for aquaculture, particularly for farming mussel (*Mytilus galloprovincialis*). Phytoplankton, primarily diatom species, account for the highest primary production in surface water layers (Miho and Witkowski, 2005; Bushati, 2013; Miho *et al*. 2013). Experts recognize phytoplankton as a reliable bioindicator of trophic status (Bushati, 2013; Kolitari *et al*.2013).

The Trophic State Index (TSI) categorizes water bodies based on their biological productivity. Data from various studies (Miho *et al*. 2013; Bushati, 2013; Çako *et al*. 2013; Kolitari *et al*. 2013; Bacu *et al*. 2022; Bacu and Zaho, 2022; Topi *et al*. 2013c; Omeri *et al*. 2024) indicate a shift in the surface waters of Butrinti Lake from oligotrophic toward mesotrophic conditions.



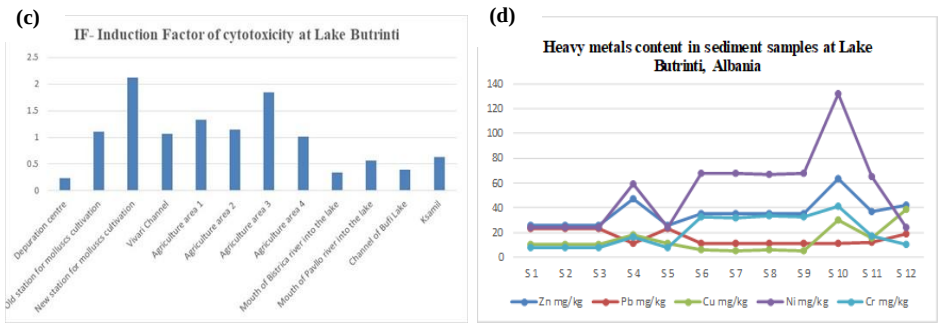


Fig. 2: ***(a)*** Factors affecting the cytotoxicity of water in the Butrinti aquatic ecosystem. ***(b)*** Trophic state of 12 stations in shallow waters at Lake Butrinti (acc to Omeri *et al*. 2024); ***(c)*** Induction Factor of cytotoxicity in waters (acc to Bacu *et al*. 2024); ***(d)*** Heavy metals content in sediment samples at Lake Butrinti (Bacu *et al*. 2024).

Algal blooms

Previous studies have indicated that Lake Butrinti provides favourable conditions for the growth of algal blooms. Consequently, dystrophic crises, including monospecific blooms, may occur and pose significant ecological risks (Miho, 1994; Bacu and Zaho, 2022). The majority of phytoplankton species belong to the *Bacillariophyceae* (diatoms) and *Dinophyceae* (dinoflagellates) (Table 9).

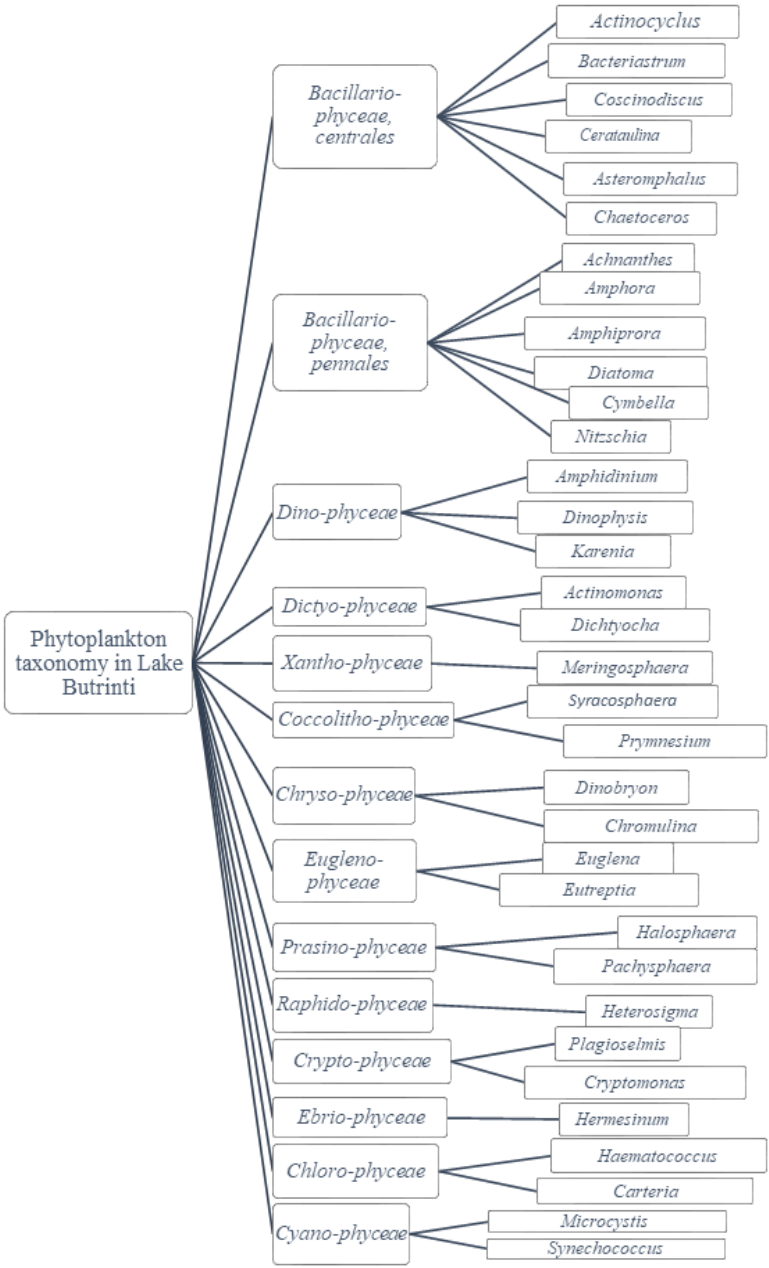


Fig. 3: Phytoplankton diversity in Lake Butrinti up to the genus level. The figure was prepared using previous data from Protić (1907), Miho (1994); Dedej (2005), and Bushati 2013.

Table 9. Algal blooms recorded in Butrinti Lake over the years.

Month	*Year*	*Species*
April	1987	*Pseudo-nitzschia* spp.
May	1987	*Prorocentrus micans*
March	1990	*Chaetoceros* spp., *Bacertiastrum* spp., *Cyclotella caspia, etc.*
March	1991	*Chaetoceros* spp. *(C. wighamii)*
January	2006	*Pseudo-nitzschia delicatissima and P. seriata*
October	2006	*Pseudo-nitzschia delicatissima and P. seriata*

Diversity of microalgae

Studies conducted between 1987 and 1991 reported 90 species of microalgae, of which 60 were diatoms (Miho, 1994; Xhulaj *et al*. 2008), followed by dinoflagellates (Miho, 1994). These species were classified into five classes: *Bacillariophyceae, Dinophyceae, Chrysophyceae, Euglenophyceae, and Cyanophyceae*, with diatoms occupying a dominant position (Miho, 1994; Peja *et al*. 1996; Xhulaj *et al*. 2008). Between 2006 and 2010, Bushati (2013) recorded approximately 460 taxa of microscopic algae across 13 classes, with diatoms and dinoflagellates being the predominant groups. Over 20 taxa have been identified as toxic, including five diatom species and 15 dinoflagellates (Miho *et al*. 2013). Bacu *et al.* (2022) reported the presence of unicellular cyanobacteria, such as *Prochlorococcus* and *Synechococcus*, emphasizing that favourable natural and anthropogenic conditions may promote their blooms, potentially posing ecological risks. Figure 3 presents a dendrogram of the phytoplankton identified to date in Lake Butrinti, classified at the genus level.

Methods used to determine phytoplankton diversity

Several analytical techniques have been employed to investigate phytoplankton diversity in Lake Butrinti, including optical, electron, and fluorescence microscopy, specific PCR assays, flow cytometry, chemotaxonomy, and catalyzed reporter deposition–fluorescence *in situ* hybridization (CARD-FISH) (Fig. 5).

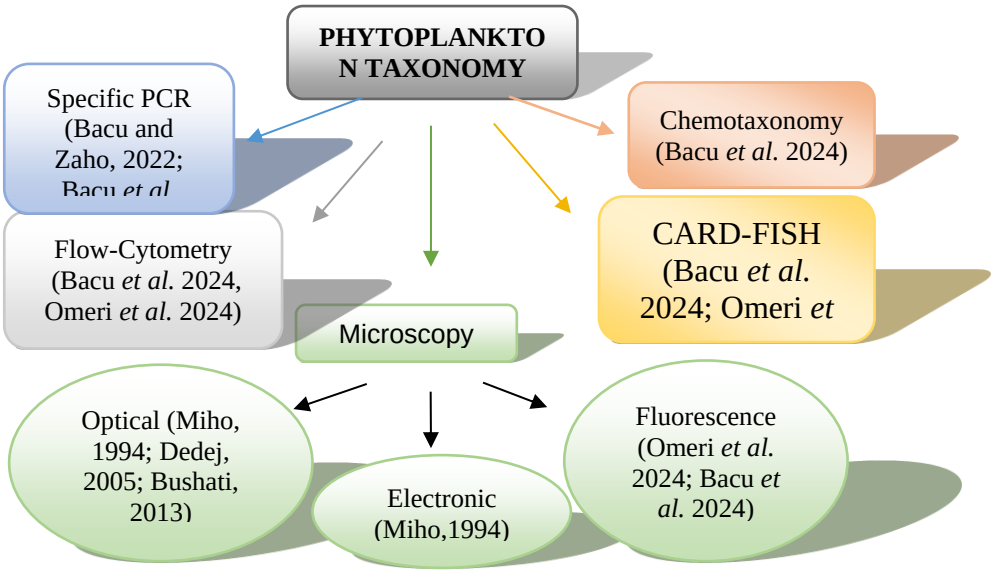

PHYTOPLANKTON TAXONOMY
Specific PCR (Bacu and Zaho, 2022; Bacu et al
Chemotaxonomy (Bacu et al. 2024)
Flow-Cytometry (Bacu et al. 2024, Omeri et al. 2024)
CARD-FISH (Bacu et al. 2024; Omeri et
Microscopy
Optical (Miho, 1994; Dedej, 2005; Bushati, 2013)
Electronic (Miho,1994)
Fluorescence (Omeri et al. 2024; Bacu et al. 2024)

Fig. 4: Techniques used to investigate phytoplankton diversity in Lake Butrinti.

Microscopy

Most phytoplankton taxonomy results for Lake Butrinti have been obtained using optical microscopy (Miho, 1994; Dedej, 2005; Bushati, 2013), with electron microscopy applied only in a few cases (Miho and Witkowski, 2005). These methods rely on the observation of morphological characteristics of phytoplankton samples, often enhanced by staining with dyes such as Lugol's solution and Methylene Blue.

DNA based methods (CARD-FISH, specific PCR, Genome Sequencing)

CARD-FISH (catalyzed reporter deposition–fluorescence *in situ* hybridization) was used to detect ribosomal RNA (rRNA), messenger RNA (mRNA), and functional genes encoded on the chromosome (Fig. 5a). This advanced FISH method is widely applied for the identification of phytoplankton and other environmental microorganisms. This technique involves designing a specific oligonucleotide probe targeting the gene or RNA of interest, which is labelled with horseradish peroxidase (HRP) enzyme. In the presence of tyramide, the signal is amplified via fluorescence, enabling the sensitive detection of the target. CARD-FISH allows for the prediction of the relative abundance of individual microbial lineages based on their rRNA gene phylogeny. It can also be combined

with direct enumeration methods, such as microscopy and flow cytometry, to estimate the absolute abundance of microscopic lineages in environmental samples (Piwosz *et al*. 2021).

Hybridized probes are typically visualized using epifluorescence microscopy, although flow cytometry is occasionally used. Although the main steps of the CARD-FISH protocol are generally consistent across studies, minor modifications may be applied, such as additional staining with DAPI (Fazi *et al*. 2007).

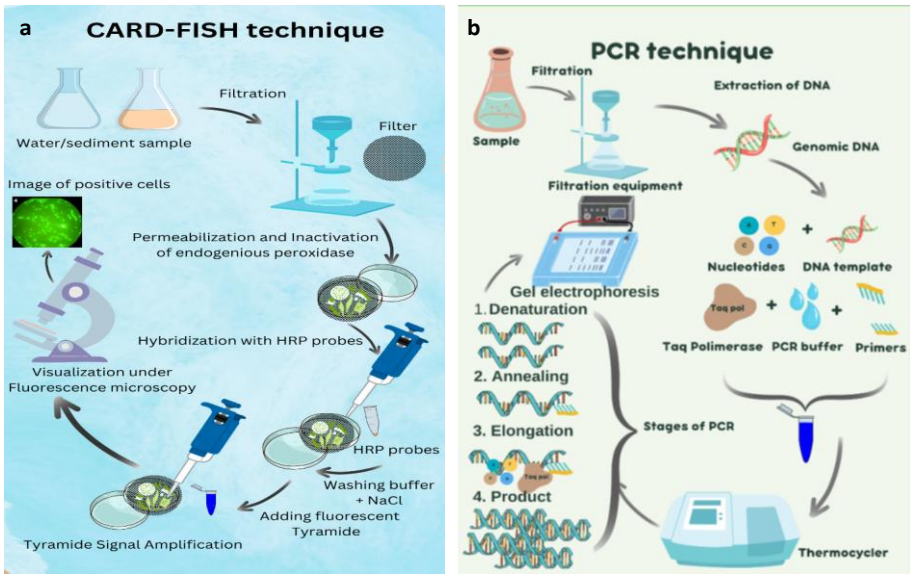


Fig. 5: ***(a)*** CARD-FISH (protocol in brief) used to investigate phytoplankton taxonomy at the class level. ***(b)*** PCR (protocol in brief) was used to determine phytoplankton genetic diversity up to the genotype/subspecies level.

Recently, CARD-FISH has been applied to assess phytoplankton biomass in Lake Butrinti (Bacu *et al*. 2024; Omeri *et al*. 2024). For the hybridization of the filter sections, different HRP-labelled oligonucleotide probes were used to target *Bacteria* (probes EUB338 I–III) and *Archaea* (probe ARCH915). The results indicated that 39.4% of the detected species belonged to *Proteobacteria*, 30% to *Bacteroidetes*, and 10% to *Cyanobacteria*. In the sediment samples, the proportion of *Cyanobacteria* fell below 1%.

Specific Polymerase Chain Reaction

PCR is one of the most commonly used molecular techniques for taxonomic analyses. It exponentially amplifies specific DNA sequences through a three-step procedure that is repeated over numerous cycles (Fig. 4b). The most frequently used genetic markers for phytoplankton taxonomy are 16S rRNA and 18S rRNA, which are universally present in bacteria, often in multiple copies or operons. The conserved function of the 16S rRNA gene over time enables more accurate phylogenetic and evolutionary analyses, and its length (~1,500 bp) is suitable for such computational analyses. The 18S rRNA gene is primarily used for phylogenetic analysis and biodiversity assessment, particularly in closely related centric diatoms (Armbrust and Galindo, 2001). The internal transcribed spacer (ITS) region is valuable for molecular systematics and population genetics; however, the presence of multiple non-identical rRNA operons can complicate comparative studies and restriction enzyme analyses (Boyer *et al*. 2001). Investigations of 16S–23S rDNA ITS diversity in Lake Butrinti (Bacu *et al*. 2022, 2024; Omeri *et al*. 2024) have revealed variations in ITS regions attributable to water pollutants. For pico-cyanobacteria, multiple ITS products were detected: ITS-a showed three sizes (1000, 900, and 300 bp), while ITS-b exhibited two sizes (550 and 350 bp).

Phytoplankton Genome Sequencing

Genome sequencing can identify subspecies and can be applied to whole genomes or specific genomic fragments (Tang *et al*. 2019). Commonly used markers include 16S rRNA (Honda *et al*. 1999; Sánchez-Baracaldo, 2015), 18S rRNA (Wang *et al*. 2014), and rbcL (Salmaso *et al*. 2022). In Lake Butrinti, DNA sequencing has been used to investigate polygenic relationships among species groups (Bacu *et al*. 2025, unpublished data) through meta-rDNA amplicon sequencing of the hypervariable V4 region of 16S rRNA, followed by taxonomic analysis.

Pigment- based method (Chemotaxonomy) for phytoplankton diversity evaluation

Chemotaxonomy uses phytoplankton pigments to investigate taxonomy, often at the class level, via CHEMTAX software (Fig. 6). Pigments are extracted from filtered water samples and serve as markers

for broad taxonomic groups (Latasa, 2007; Develi *et al*. 2012; Matek *et al*. 2023; Peltomaa *et al*. 2023). A simplified CHEMTAX approach applied to Lake Butrinti samples revealed changes in phytoplankton community composition associated with climate change (Bacu *et al*. 2024).

Table 10. An example of chemotaxonomy based evaluation of the main taxa at Butrinti Lake considering the simplified CHEMTAX approach

Stations	klb/kla	klc/kla	Taxonomy
B1-0	0.030481377	0.083801827	*Haptophyceae, Bacillariophyceae, Chrysophyceae, Proclorophyceae*
B1-8	3.932308084	0.00687638	*Proclorophyceae, Prasinophyces*
B2-0	0.855646268	0.711292537	*Chrysophyceae, Haptophyceae, Proclorophyceae, Prasinophyceae*
B2-22	1.791247622	0.271359971	*Dinoflagjelate, Haptophyceae, Chrysophyceae, Diatoms, Proclorophyceae, Prasinophyces*
B3-0	1.350116587	2.318815951	*Prasinophyces*
B3-19	0.359337349	0.428915663	*Dinoflagjelate, Prasinophyceae, Chrysophyceae, Proclorophyces, Euglenophyceae, Prasinophyceae*

The method was applied to samples collected from stations B1–B3, with sampling performed at both the surface and maximum depth: B1-0 (Station 1, depth 0 m), B1-8 (Station 1, depth 8 m), B2-0 (Station 2, depth 0 m), B2-22 (Station 2, depth 22 m), B3-0 (Station 3, depth 0 m), and B3-19 (Station 3, depth 19 m), as reported by Bacu *et al*. (2024).

Each method described above has specific advantages and limitations (Table 11), which should be considered in relation to the research objectives, whether focusing on phytoplankton taxonomy, water quality

and trophic state assessment, phytoplankton abundance, community composition, genetic diversity, or the evaluation of pollutant effects on cytotoxicity and subspecies-level genetic variation.

Table 11. Summary of main uses and limitations of methods for water quality and phytoplankton diversity evaluation, applied at Lake Butrinti

Methods	**Purpose of use**	**Limitations**	**References**
Physical-chemical parameters of waters and sediments	Determine the quality of waters according to accepted standards.	Cannot be used to investigate phytoplankton community composition	USEPA, 1991
Biomarkers	Determine Trophic State/Quality of waters according to accepted standards.	Cannot be used to interpret the chemical composition of waters	USEPA, 1991
Biosensors (single-cell fiber-optics, luminescent bacteria, etc)	Evaluate the cytotoxicity imposed by different pollutants in water and sediments.	Higher sensitivity is required in complex wastewater environments, long-term stability, and regulatory barriers must be addressed.	Fdez-Sanromán *et al*. 2025
Microscopy (optical, electronic, fluorescence)	Identification to species level, evaluation of abundance.	Time-consuming; The reliability of identification depends on the skills of the identifier. Morphologically similar species cannot be distinguished from each other.	Peltomaa *et al*. 2023
Flow cytometry	Counting, analysis, and identification of phytoplankton species and groups.	Standard flow cytometry (FCM) is restricted to a low taxonomic resolution, making it unsuitable for identifying indicator species. Imaging flow cytometers (IFC) can overcome these limitations.	Dunker *et al* 2019
Chemotaxonomy	Monitoring phytoplankton communities;	Limitations due to overlapping profiles of some biomarkers.	Peltomaa *et al*. 2023

	Taxonomy at the class level identification;		
CARD-FISH	Estimates of relative abundance (percent contribution to total eukaryotic numbers) of individual microbial lineages defined by their rRNA gene phylogeny; It can be separately optimized for each target group (probe), which is not possible for PCR with primers that target many different templates.	Imperfect probe coverage and specificity; poor detection of low abundance or inactive community members, and difficulties in counting aggregated cells, autofluorescence of chloroplasts might also interfere with probe signals; and tyramides sometimes bind unspecifically.	Kubota,2013 Piwosz *et al*. 2021
Amplicon and shotgun Sequencing & Metabarcoding	Species identification based on the similarity or divergence of the molecular sequence of an unknown organism to a vouchered reference sequence in the database.	Intragenomic variation with many copies of individual genes causes overestimation of cell numbers and diversity; An annotated barcode library and choice of barcode and primers are essential.	Alemzadeh. *et al*. 2014 Gelis *et al*. 2024

3. CONCLUSIONS

Lake Butrinti represents a dynamic aquatic ecosystem shaped by the combined influences of freshwater inflow, saline intrusion from the Ionian Sea, and contributions from Lake Bufi. Over time, it has exhibited a trend toward mesotrophic conditions, underscoring the need for continuous monitoring of anthropogenic activities and climate change, the two primary influencing factors of water quality. Strict regulation of fertilizer and pesticide use in agriculture, proper treatment of urban wastewater, controlled diversion of the Bistrica River during summer, and ongoing maintenance of the Butrinti Canal (enlargement and deepening) are essential to mitigate potential ecological and socio-economic impacts, such as dystrophic crises, declines in fishing and mussel production, loss of

biodiversity, degradation of ecological quality, and reduced environmental aesthetic value. Furthermore, there is a clear need for more comprehensive and systematic research employing both conventional and advanced methods to monitor ecological changes, mitigate negative effects, and safeguard the integrity and economic value of the lake. Emphasis should be placed on the specific strengths of each analytical approach—physicochemical, biological, DNA-based, and biosensor methods—and their integration into a unified monitoring framework to ensure food safety. Integrating these methodologies into the monitoring program of the Albanian National Agency of Environment would strengthen long-term ecological assessment and provide a robust evidence base to guide environmental policy implementation in Albania.

DECLARATIONS

Data accessibility: There are no databases to be made public**.**

Declaration of AI use: There has been no use of AI when writing the actual paper.

Author's contribution:

Xh.O.: writing-original, editing draft: A.B.: conceptualization, validation, writing-review and editing.

Authors approval

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration: The authors declare there are no conflicts of interest.

Funding: The authors declare no funding was received when writing this paper.

Ethics: This work did not require ethical approval from individuals prior to publication.

REFERENCES:

Alemzadeh E, Haddad R, Ahmadi AR. 2014. Phytoplanktons and DNA barcoding: Characterization and molecular analysis of phytoplanktons on the Persian Gulf. *Iranian Journal of Microbiology*, **6(4):** 296-302. PMID: 25802715; PMCID: PMC4367948.

Anonymous. 2010. Plani i Menaxhimit (version përfundimtar i rishikuar) te PK Butrint_ASPBM. SHRSHGJSH: 1-229(http://www.moe.gov.al/upload/zona%20te%20mbrojtura/Plani_MenaxhimitPK%20Butrint_2010_ALB.pdf).

Ariztegui D, Anselmetti F, Robbiani J-M, Bernasconi S, Brati E, Gilli A, Lehmann MF. 2010. Natural and human-induced environmental change in southern Albania for the last 300years — Constraints from the Lake Butrint sedimentary record. *Global and Planetary Change*, **71 (3–4):** 183-192. DOI:10.1016/j.gloplacha.2009.11.016.

Armbrust V, Galindo M. 2001. Rapid evolution of a sexual reproduction gene in centric diatoms of the genus Thalassiosira. *Applied and Environmental Microbiology*, **67 (8):** 3501-3513. https://doi.org/10.1128/AEM.67.8.3501-3513.2001.

Bacu A, Omeri Xh, Sadikaj R, Puto K. 2024. An integrated approach based on the use of monitoring and research protocols for water quality assessment in Albania. *Second International Conference on Public Health and Environment, Water and Air Quality Control: Protecting Human Health and Ensuring Sustainable Developmen*t. Page 51.

Bacu A, Suli K, Zaho I. 2022. Diversity of picocyanobacteria and quality of waters at anoxic lake butrinti. 4rth International Conference on Applied Biotechnology, Ioannina, Greece. *Book of Proceedings*, pg 21-28, ISBN: 978-960-233-276-4.

Bacu A, Zaho I. 2022. Determining environmental factors that may influence cyanobacteria bloom formation in Lake Butrinti, Albania. *Journal of Environmental Protection and Ecology*, **23 (1):** 1-8. https://scibulcom.net/en/journal/1311-5065/issue/2022-23-1/.

Bego F, Malltezi J. 2011. Ecotourism opportunities and challenges in Butrint, Albania, a unique UNESCO and Ramsar site. *Proceedings of the MCRR3*. Italy. https://doi.org/10.2112/SI61-001.9

Bond GP, Martin J. 2005. *Microtox, encyclopedia of toxicology* (Second Edition), Elsevier, Pages 110-111, ISBN 9780123694003, https://doi.org/10.1016/B0-12-369400-0/00630-X.

Boyd EC. 2001. Through its responsible aquaculture program, GAA has recommended a range of water quality standards for shrimp farm effluents. Part of a series, this article addresses total phosphorus. https://www.globalseafood.org/advocate/water-quality-standards-total-phosphorus/#.

Boyer SL, Flechtner VR, Johansen JR. 2001. Is the 16S–23S rRNA internal transcribed spacer region a good tool for use in molecular systematics and population genetics? A case study in cyanobacteria. *Molecular Biology and Evolution*, **18 (6):** 1057–1069. https://doi.org/10.1093/oxfordjournals.molbev.a003877.

Bushati M. 2013. Studim mbi prodhimtarine paresore dhe sigurine shendetesore te saj ne liqenin e Butrintit (Sarande). PhD Thesis, FNS, UT, Tirana: 195 pp. (in Albanian). https://unitir.edu.al/doktoratura-merjem-bushati-fakulteti-i-shkencave-i-natyrore-departamenti-i-biologjise/.

Çako V, Baci S, Shena M, 2013. Water turbidity as one of the trophic state indices in Butrinti Lake. *Journal of Water Resource and Protection*, **5 (12):** 1144-1148. 10.4236/jwarp.2013.512120.

Dedej Z, Bino T. 2003. Information Sheet on Ramsar wetlands (RIS) categories approved by recommendation 4.7, as amended by Resolution VIII.13 of the Conference of the Contracting. Parties. https://www.ramsar.org/sites/default/files/documents/pdf/res/key_res_viii_13_e.pdf.

Dedej Z. 2005. Alga mikroskopike nga mjedise bregdetare shqiptare. PhD Thesis, FNS, UT, Tirana: 106 pp. (in Albanian).

Develi EE, Berthon JF, Canuti E, Slabakova N, Moncheva S, Shtereva G, Dzhurova B. 2012. Phytoplankton taxonomy based on CHEMTAX and microscopy in the northwestern Black Sea. *Journal of Marine Systems*. **94:** 18-32. https://doi.org/10.1016/j.jmarsys.2011.10.005.

Dunker S. 2019. Hidden secrets behind dots: Improved phytoplankton taxonomic resolution using high-throughput imaging flow cytometry. *Cytometry A*, **95(8):**854-868. doi: 10.1002/cyto.a.23870. Epub 2019 Aug 6. PMID: 31385646.

European Commission 2008: *Framework Water Directive on environmental quality standard.* 39. https://environment.ec.europa.eu/topics/water/water-framework-directive_en.

European Union. 2013. Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. Official Journal of the European Union, L 226, 1–17. https://data.europa.eu/eli/dir/2013/39/oj (EQS 2013/39/EU)

Fazi S, Amalfitano S, Pizzetti I, Pernthaler J. 2007. Efficiency of fluorescence in situ hybridization for bacterial cell identification in temporary river sediments with contrasting water content. *Systematic and Applied Microbiology*, **30 (6):** 463-470, ISSN 0723-2020, https://doi.org/10.1016/j.syapm.2007.03.003.

Fdez-Sanromán A, Bernárdez-Rodas N, Rosales E, Pazos M, González-Romero E, Sanromán MÁ. 2025. Biosensor technologies for water quality: Detection of emerging contaminants and pathogens. *Biosensors*, 15 (3): 189. https://doi.org/10.3390/bios15030189 .

Gabeira H. 2023. Study of surface water quality and microsystem limnology. *Journal of Aquaculture Research & Development*, **14 (2):** 729. DOI: 10.35248/2155-9546.23.14.729. ISSN:2155-9546.

Gelis MRN, Canino A, Bouchez A, Domaizon I, Laplace-Treyture C, Rimet F, Alric B. 2024. Assessing the relevance of DNA metabarcoding compared to morphological identification for lake phytoplankton monitoring. *Science of the total environment*, **914**: 169774, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2023.169774.

Guri S, Guri M, Muhedini B, Naska M, Vidos J, Guri, A, Kodra K, Kalemi E. 2014. *Raport për qendrën e grumbullimit dhe depurimit të midhjes, Manastir/Sarandë.* Agjencia Kombëtare e Mjedisit. https://akm.gov.al/wp-content/uploads/2020/07/vnm-paraprak-midhje.pdf.

Heywood D. 2017. *Environmental and social management plan (ESMP) – Xarra – Final Version.* Ministry of Agriculture, Rural Development and Water Administration. (PROJECT ID NO: P121186 – RFP NO: MAFCP/CS/007 – LOAN NO. 8211 – AL). https://bujqesia.gov.al/wp-content/uploads/2017/12/ESMP_Xarra_ID_-_Ver_E_after_PC.pdf.

Honda D, Yokota A, Sugiyama J. 1999. Detection of Seven Major Evolutionary Lineages in Cyanobacteria Based on the 16S rRNA Gene Sequence Analysis with new sequences of five marine *Synechococcus* strains. *Journal of Molecular Evolution*, **48(6):**723-39. doi: 10.1007/pl00006517.

Hounslow MW, Chepstow-Lusty A. 2004. Record of soil loss from Butrint, southern Albania, using mineral magnetism indicators and charcoal (AD 450 to 1200). *The Holocene*, **14(3):** 321–333. https://doi.org/10.1191/0959683604hl705rp.

Këshilli i Ministrave i Republikës së Shqipërisë. 2002. Vendimi nr. 531, datë 13.11.2002, Për shpalljen e Lagunës së Butrintit si zonë të mbrojtur Ramsar. *Fletorja Zyrtare e Republikës së Shqipërisë,* 2002, 155, 4370–4372. (DCM 531/2002)

Kolitari J, Gjyli L, Mukli L, Gjyli S, Vukaj J. 2013. Distribution of Chlorophyll a in Lagoon of Butrint waters comparing with environment factors (Albania). *Albanian journal agricultural science*, 87-93. ISSN: 2218-2020.

Kubota K. 2013. Minireview CARD-FISH for environmental microorganisms: technical advancement and future applications. *Microbes and Environment*, **28(1):** 3–12. doi: 10.1264/jsme2.ME12107.

Latasa M. 2007. Improving estimations of phytoplankton class abundances using CHEMTAX. *Marine ecology progress series*, **329**: 13–21, 2007. DOI: https://doi.org/10.3354/meps329013.

Matek A, Bosak S, Šupraha L, Neeley A, Višić H, Cetinić I, Ljubešić Z. 2023. Phytoplankton diversity and chemotaxonomy in contrasting North Pacific ecosystems. *PeerJ11*: e14501 http://doi.org/10.7717/peerj.14501 .

Miho A, Kashta L, Beqiraj S 2013. *Between the land and the sea - ecoguide to discover the transitional waters of Albania.* Julvin 2, Tiranë. 462 pp. ISBN 978-9928-137-27-2. https://fshn.edu.al/info/publikime-shkencore; https://www.researchgate.net/publication/261323171_Between_the_Land_and_the_Sea_-_Ecoguide_to_discover_the_transitional_waters_of_Albania.

Miho A, Witkowski A. 2005. Diatom (Bacillariophyta) Flora of Albania Coastal Wetlands Taxonomy and Ecology: A Review. *Proceedings of the California Academy of Sciences*, **56(12):** 129-145, 1 figure, 2 plates, Appendix. https://archive.org/details/biostor-249129/mode/2up;https://ia801005.us.archive.org/23/items/biostor-249129/biostor-249129.pdf.

Miho A. 1994. Qualitative and quantitative data on the phytoplankton of Butrinti lake (Saranda). PhD Thesis, FNS, UT, Tirana: 145 pp. (in Albanian).

Moisiu L, Panagiotopoulos I, Durmishi C, Kapsimalis V, Anagnostou CL. 2016. The anoxic Butrint Lagoon, SW Albania. *Environmental Earth Science*, **75**:1443. https://doi.org/10.1007/s12665-016-6259-0.

Morellón, M, Sinopoli, G, Izdebski, A, Sadori, L, Anselmetti, F, Hodges, R, Regattieri, E, Wagner, B, Brushulli, B, and Ariztegui, D. 2016. Environment, Climate and Society in Roman and Byzantine Butrint. *Late Antique Archaeology*, **12 (1):** 3-17. https://doi.org/10.1163/22134522-12340063.

National Environmental Agency (NEA). 2015. Albania state of the environment report 2014. Tirana: National Environmental Agency. Retrieved from https://akm.gov.al/wp-content/uploads/2022/03/raporti-2014-i-perfunduar.pdf

Nuro A, Marku E. 2011. Organochlorine pesticides residues for some aquatic systems in Albania. In M. Stoytcheva, *Pesticides - Formulations, Effects, Fate*.10.5772/13972. DOI:10.5772/13972.

Omeri Xh, Bacu A, Canameti S, Fazi S. 2024. Physical-biological indicators and single-cell toxicity sensing as a new monitoring model applied at Lake Butrint, Albania. *Proceedings of the Bio-X*, Zagreb, Croatia. Pg. 90-100; ISBN: 978-953-6076-76-5.

Omeri Xh, Bacu A, Zaho I, Canameti S. 2025. Determining of phytoplankton biomass and composition at Butrinti Lake-An overview of methods efficiency. European Biotechnology Congress 2024. *The EuroBiotech Journal*, in press.

Osmani Miri F, Peja N. 2012. Preliminary data on protozoan community in Butrinti lagoon (Southern Albania). International Conference on Marine and Coastal Ecosystems (MarCoastEcos 2012): increasing knowledge for a sustainable conservation and integrated management, 26 – 29 April 2012, Tirana. http://marcoastecos2012.al/.

Pano N. 1984. *Hidrologjia e Shqipërisë*. Tirana: Akademia e Shkencave, Caption X, 397-405.

Pano N. 2015. *Pasuritë Ujore të Shqipërisë*, 2nd edition, Akademia e Shkencave: Tiranë; 349-350 and 386-396. ISBN 978-99956-10-87-6.

Peja N, Vaso A, Miho A, Rakaj N, Crivelli AJ. 1996. Characteristic of Albanian lagoons and their fisheries. *Fisheries Research***, 27 (4):** 215-225. https://doi.org/10.1016/0165-7836(95)00467-X.

Peltomaa E, Asikainen H, Blomster J, Pakkanen H, Rigaud C, Salmi P, Taipale S. 2023. Phytoplankton group identification with chemotaxonomic biomarkers: In combination they do better. *Phytochemistry*, 209: 113624, ISSN 00319422. https://doi.org/10.1016/j.phytochem.2023.113624.

Piwosz K, Mukherjee I, Salcher MM, Grujčić V, Šimek K. 2021. CARD-FISH in the sequencing era: Opening a new universe of

Protistan ecology. *Frontiers in Microbiology*, 12 – 2021. https://doi.org/10.3389/fmicb.2021.640066.

Protić G. 1907. Beitrag zur Kenntnis der Algenflora Albaniens. Wissenschaftliche Mitteilungen aus Bosnien und Herzegovina 10:611–621. https://www.zobodat.at/pdf/Wiss-Mitt-Bosnien-Hercegovina_10_1907_0611-0621.pdf.

Salmaso N, Vasselon V, Rimet F, Vautier M, Elersek T, Boscaini A, Donati C, Moretto M, Pindo M, Riccioni G, Stefani E, Capelli C, Lepori F, Kurmayer R, Mischke U, Klemenčič AK, Novak, K, Greco C, Franzini G, Fusato G, Giacomazzi F, Lea A, Menegon S, Zampieri Ch, Macor A, Virgilio D, Zanut E, Zorza R, Buzzi F, Domaizon I. 2022. DNA sequence and taxonomic gap analyses to quantify the coverage of aquatic cyanobacteria and eukaryotic microalgae in reference databases: Results of a survey in the Alpine region. *Science of The Total Environment*, **834**: 155175, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2022.155175.

Sánchez-Baracaldo P. 2015. Origin of marine planktonic cyanobacteria. *Scientific Reports*, **5:** 17418. https://doi.org/10.1038/srep17418.

T**ang, J, Du, LM, Liang, YM, Daroch M. 2019**. Complete genome sequence and comparative analysis of Synechococcus sp. CS-601 (SynAce01), a cold-adapted cyanobacterium from an Oligotrophic Antarctic Habitat. *International Journal of Molecular Sciences*, **20(1):**152. DOI:10.3390/ijms20010152.

Topi T, Bani A, Malltezi J, Sulce S. 2012. Heavy metals in soil, sediments, mussels, and water from Butrinti Lagoon (Albania). *Fresenius Environmental Bulletin*. **21**. 3042-3051. WOS:000312412700007.

Topi T, Bani A, Sulce S. 2013a. Physico chemical characteristics and heavy metal contents of water from Butrinti lagoon, Albania. *Albanian Journal of Agricultural Sciences*, 321-326. ISSN: 2218-2020.

Topi T, Bani A, Sulce S. 2013b. Risk assessment of Butrinti lagoon: Spatial and temporal distribution of heavy metal s in different pools. *Albanian Journal of Agricultural Sciences*, **12 (1):** 25-3. ISSN: 2218-2020.

Topi M, Mersinaj K, Saliaj O. 2013c. Preliminary report for the key biodiversity area of Butrinti. protection and preservation of natural environment in Albania, PPNEA. Preliminary-Report-KBA-Butrinti.pdf

UNESCO. (n.d.). Retrieved from UNESCO, World Heritage Convention: https://whc.unesco.org/en/list/570/

United States Environmental Protection Agency. (USEPA). 1991. Methods for evaluating the toxicity of chemical pollutants. U.S. Environmental Protection Agency, Office of Research and Development. https://www.epa.gov/wqs-tech/what-are-water-quality-standards

Velaj T. 2015. Geological Setting and Coastal-marine ecosystem of Butrinti region, Albania. *Journal of Petroleum and Environmental Biotechnology*, **6 (6):** 252. DOI: 10.4172/2157-7463.1000252.

Wang Y, Tian RM, Gao ZM, Bougouffa S, Qian PY. 2014. Optimal eukaryotic 18S and universal16S/18S ribosomal RNA primers and their application in a study of symbiosis. *PLoS One*; **9(3):** e90053. DOI:10.1371/journal.pone.0090053.

Woutersen M, Belkin, SH, Brouwer A, Wezel A, Heringa M. 2010. Are luminescent bacteria suitable for online detection and monitoring of toxic compounds in drinking water and its sources? *Analytical and Bioanalitycal Chemistry*, **400 (4):**915-29. doi: 10.1007/s00216-010-4372-6.

Xhulaj S, Bushati M, Miho A. 2008. Overview on phytoplankton of Albanian lagoons. The III Congress of Ecologists of the Republic of Macedonia with International Participation, 8. Skopje. https://www.mes.org.mk/PDFs/3rd%20Congress%20Proceedings/07_Skerdilaid%20Xhulaj,%20Merjem%20Bushat,%20Aleko%20Miho.pdf.

Zotaj A. 2010. Butrinti national park management plan, *Integrated Coastal zone management and clean-up project*. Project ID: PO86807, https://www.academia.edu/28471210/Butrint_MP_English_version_September_2010.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/ which permits unrestricted use, provided the original author and source are credited.

AN INTEGRATED APPROACH BASED ON THE USE OF MONITORING AND RESEARCH PROTOCOLS FOR WATER QUALITY ASSESSMENT IN ALBANIA

Ariola BACU
Department of Biotechnology, Faculty of Natural Sciences, University of Tirana, Albania
Centre for Biotechnology and Genetics, Albanian Academy of Sciences

Xhensila OMERI, Rigerta SADIKAJ and Klementina PUTO
Department of Biotechnology, Faculty of Natural Sciences, University of Tirana, Albania
Author for correspondence: ariola.bacu@fshn.edu.al

AB, 0000-0001-6390-2632 ; XhO, orcid.org/0009-0005-3937-6245; RS, 0000-0002-5311-6205; KP, 0000-0002-8957-7724
https://doi.org/10.67537/ajnts3001008

ABSTRACT

Water quality is one of the most critical environmental issues in the context of climate change, as are the methodologies applied to evaluate it. In Albania, standardized methods have traditionally been employed for the measurement of physical-chemical indicators, trophic state (ISO 17025), and bacteriological indicators (ISO16649-3:2015). However, the use of any these methodologies in isolation does not provide a comprehensive understanding of the condition of fresh, marine or lagoon ecosystems. The present study discusses an integrated approach to water quality evaluation that combines standardized monitoring protocols with advanced biotechnological research methods. These include CARD-FISH, fluorescence microscopy, and flow cytometry for biomass quantification; factorial analysis of bacterial DNA in relation to environmental parameters; assessment of phytoplankton genetic diversity; and chemotaxonomic analyses. In addition, the application of novel standardized monitoring protocols employing bioreporter bacteria for cytotoxicity testing is presented. The results obtained from the implementation of these methodologies in Albania indicate that an integrated, multidisciplinary approach provides a more comprehensive and reliable framework for the assessment of water quality.

Keywords: single-cell biosensors, CARD-FISH, chemotaxonomy, Flow cytometry, trophic state

1. INTRODUCTION

Monitoring of water quality in rivers, lakes, lagoons, and coastal areas has long been part of the services provided by the Department of Biotechnology, University of Tirana, to the National Agency of Environment of Albania until 2015. Since then, the department's capacities and expertise in this field have been primarily devoted to research activities. However, according to the Albanian Law on Higher Education Institutions, university departments are expected to engage in teaching, research, and third-party services. The latter plays a crucial role for both universities and external stakeholders, including public and private institutions and businesses.

ISO protocols and/or research methodologies have been applied to address the following parameters: i) physical-chemical indicators, ii) pathogenic bacteria using indicator species, iii) trophic state (Carlson Index), iv) cytotoxicity sensing, v) phytoplankton community composition, vi) use of phytoplankton DNA as a biomarker, vii) factorial analysis of DNA / environmental parameters, viii) pico-cyanobacteria PCR-based subspecies diversity exploration, ix) invasive species investigation & Monitoring conditions for algal blooms, and x) CARD-FISH and fluorescence imaging.

The results obtained from the application of these methodologies demonstrate that the use of ISO protocols or research-based methods alone does not provide a complete understanding of the condition of freshwater, marine, or lagoon ecosystems. Therefore, this study proposes an integrated approach that combines both standardized and research-oriented protocols for a more comprehensive evaluation of water quality. This approach enables simultaneous assessment of the chemical composition of water, the presence of pathogenic bacteria of public health importance, trophic state (reflecting the ecological health of aquatic environments), and the impacts of various pollutants (e.g., heavy metals, pesticides, herbicides) on aquatic organisms. Furthermore, it allows the examination of environmental influences on phytoplankton community structure and biomass.

The implementation of this integrated methodology in Albanian waters supports the conclusion that such an approach not only provides an accurate evaluation of the current state of aquatic ecosystems but also enhances understanding of microbial community interactions with environmental and anthropogenic factors—thus contributing to predictive modeling of future ecological conditions.

2. MATERIALS AND METHODS

The data presented in this study were obtained and analyzed using the following methodologies, applied to several major Albanian aquatic ecosystems, including Lake Butrint, Lake Ohrid, and Lake Shkodra. All procedures were conducted within research projects undertaken at the Department of Biotechnology, University of Tirana.

2.1. Physicochemical analysis of water quality

Water quality was assessed using physicochemical indicators. Temperature, pressure, dissolved oxygen, electrical conductivity, salinity, and pH were measured in situ. The remaining parameters were analyzed in the laboratory using standardized protocols, as listed below:

1. Chemical oxygen demand (COD): ISO 15705:2002
2. Biochemical oxygen demand (BOD_5): EPA Method 5210
3. Total suspended solids (TSS): S SHEN 872:2005
4. Total dissolved solids (TDS): APHA/AWWA 2540C/2022
5. Phosphates: APHA/AWWA 4500-P E/2022
6. Nitrates: ISO 7890-3:1988
7. Nitrites: ISO 6777:1984
8. Ammonium: APHA/AWWA 4500-NH_3 F/2022
9. Total nitrogen: ISO 11905-1:1997
10. Total phosphorus: ISO 6878:2004
11. Major and trace elements (K, Na, Ca, Mg, Cd, Cr, Cu, Fe, Mn, Ni, Pb, Zn): APHA/AWWA 3111B/2022

2.2. Microbiological analysis

The presence of pathogenic bacteria was investigated using indicator species according to ISO 16649-3:2015.

2.2. Trophic state determination

The trophic state of each water body was calculated using the Carlson Trophic State Index (Carlson, 1977).

2.4. Cytotoxicity sensing

Cytotoxicity was evaluated using *Vibrio fischeri* bioluminescent bacteria following ISO 11348-3:2007.

2.5. Phytoplankton community analysis

Phytoplankton community composition was determined through chemotaxonomic methods following Mackey et al. (1996). Factorial analysis was performed to explore correlations between phytoplankton DNA characteristics and environmental parameters.

2.6. Molecular and cytometric analyses

Subspecies diversity of pico-cyanobacteria was explored using the 16S rDNA ITS region. CARD-FISH and fluorescence imaging were employed for phytoplankton biomass evaluation, complemented by flow cytometry to quantify biomass and characterize cell populations.

3. RESULTS and DISCUSSIONS

- **Water quality evaluation based on physicochemical indicators: Case study of Lake Butrint (2023)**

Data on the physicochemical characteristics of Lake Butrint have been extensively published and reviewed due to the lake's environmental, economic, and cultural importance (Bacu *et al.* 2022). In this study, we analyzed a comprehensive set of parameters, including pH, temperature, electrical conductivity, salinity, total suspended solids (TSS), total dissolved solids (TDS), total organic carbon (TOC), chemical oxygen demand (COD), biochemical oxygen demand (BOD), and a range of elements (P, N, Zn, Fe, Al, Cr, As, Cd, Hg, Pb, Cu, and Ni), to assess the water quality of Lake Butrint during 2023. The assessment was conducted in accordance with the Water Framework Directive (75/440/EEC) of the European Union and Council Decision No. 379, dated 25 May 2016, approving the regulation "Drinking Water Quality." In addition, reference was made to the standards for urban wastewater discharges established under Albanian legislation (VKM No. 177, dated 31 March 2005). Based on these criteria, Lake Butrint was classified as being in a mesotrophic to eutrophic state. Seasonal fluctuations in the measured parameters were analyzed, and regression curves with corresponding equations were generated to describe variations over time. Depth profiles of selected physicochemical parameters (Fig. 1) indicate the presence of anoxic conditions at depths of 8 m and below, reflecting reduced oxygen availability and potential stratification effects within the water column.

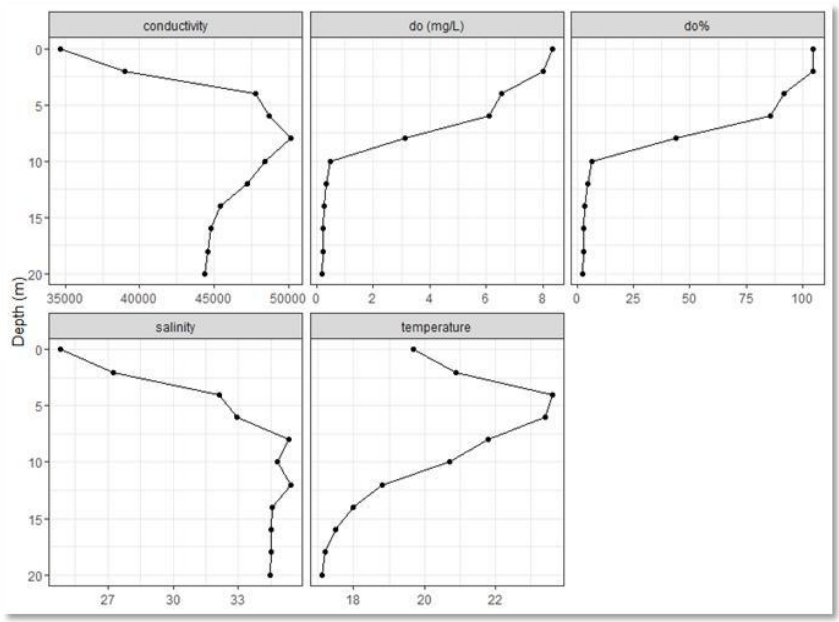


Fig.2. Depth profiles of some Physical-chemical parameters measured at Butrinti Lake.

- **Investigation of pathogenic bacteria using indicator species: Case study of Lake Butrint**

The presence of pathogenic bacteria in Lake Butrint was investigated through the monitoring of *faecal coliforms* and *faecal streptococci* during the period 2021–2022. The results were evaluated in accordance with European Union Directive 75/440/EEC on the quality required of surface waters intended for the abstraction of drinking water, as well as with the WHO/UNEP and EU guidelines for surface water classification based on the 90th and 95th percentile values.

These bacterial indicators were selected due to their reliability in reflecting recent faecal contamination and their wide use as standard parameters in microbiological water quality assessment. The observed concentrations of faecal indicator bacteria varied seasonally, with higher counts typically recorded during warmer months, reflecting increased microbial activity and potential runoff from surrounding anthropogenic sources.

Table 1. *Faecal coliform* presence at Butrinti Lake, 2022

	November 2021	December 2021	January 2022	March 2022	April 2022	May 2022	June 2022	November 2022
Station 1	11000	900	1500	<3	400	700	1500	2300
Station 2	11000	<3	<3	700	400	700	400	4300
Station 3	11000	900	<3	<3	<3	1500	400	15000

In November 2021 and 2022, the microbial load of *Faecal coliforms* at all three monitoring stations exceeded the limit value established by the European Union (2,000 bacteria per 100 mL of water).

Table 2. *Faecal streptococci* presence at Butrinti Lake, 2022

Fecal streptococci	November	December	January	March	April	May	June
Station 1 "Mussel Gathering Ponit"	<3	1100	1500	<3	<3	<3	<3
Station 2 "In the city of Ksamil"	<3	1100	430	<3	<3	<3	<3
Station 3 "Butrinti National Park"	<3	2800	2400	400	<3	<3	<3

Similarly, during December 2021 and January 2022**,** the *Faecal streptococci* load at Station 3 surpassed the EU threshold of 2,000 bacteria per 100 mL**.**

- The presence of *Escherichia coli* in mussels was also monitored over a two-month period (June–July 2022) following the ISO/TS 16649-3:2015 protocol

Table 3. *E. coli* presence in mussels at Butrinti Lake, 2022.

	Dilutions		Station I			Station II			Station III
1-15 June	MPN		**330**			**460**			**<18**
15-30 June	MPN		**<18**			**<18**			**45**
1-15 July	MPN		**<18**			**<18**			**61**
15-31 July	MPN		**<18**			**18**			**<18**

As shown in Table 3, Stations 1 and 2 exhibited a high microbial load of *Escherichia coli* during the first weeks of June 2022, exceeding acceptable microbiological limits for surface waters.

• Cytotoxicity sensing using bio-reporting bacteria- at Vlora, Albania- 2023

Luminescent bioreporter bacteria were employed to assess the cumulative toxic effects of various pollutant categories present in water and sediment samples on the native microbiota. Measurements were performed in November 2023 at eight sampling stations in the Bay of Vlora, Albania. The Luminescence Induction Factors (IF), expressing either stimulation (S) or inhibition (I) of bacterial luminescence, ranged from 0.39–3.82 S to 0.51 I (Fig. 2). Stations 1, 2, 4, 6, and 8 exhibited IF values above 1.5 S, indicating a significant level of toxicity, whereas Stations 5 and 7 showed values between 0.5 and 1.0 S, representing borderline or moderate toxicity levels.

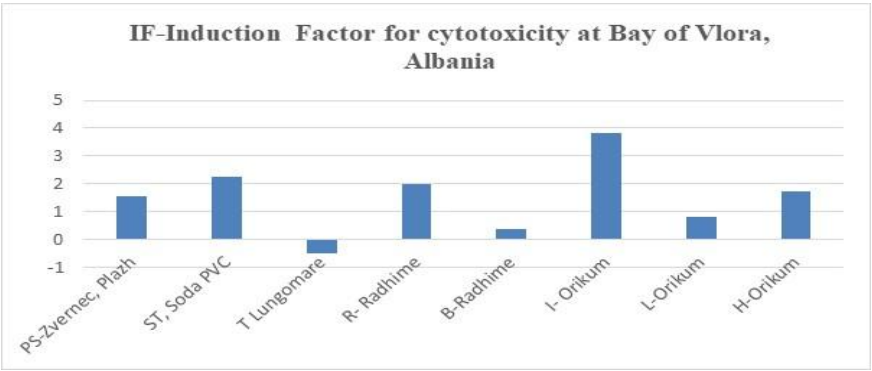


Fig. 2. Variation of cytotoxicity in Vlora Bay based on bioluminescent bio-reporting bacteria (November 2023).

- Trophic State seasonal variation- Case of Butrinti Lake

The Trophic State Index (TSI)**,** as proposed by *Carlson (1977),* utilizes algal biomass to classify the trophic condition of aquatic systems.

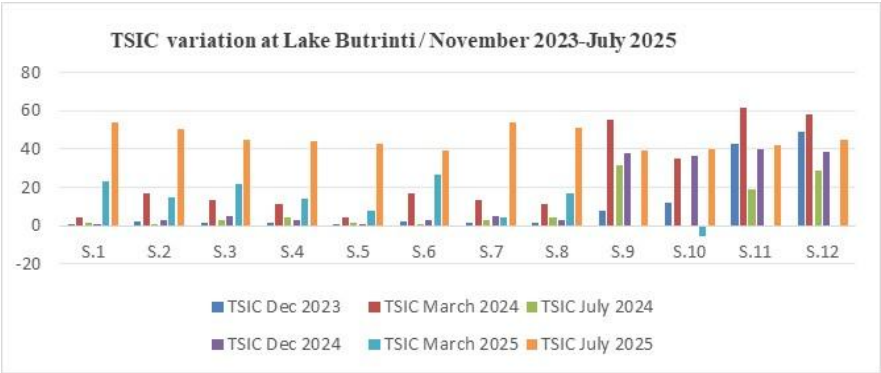


Fig. 3. Trophic State (Carlson Index based on chlorophyll a) at Lake Butrinti shallow waters from November 2023 to July 2025.

Results indicate both spatial and temporal variations in the trophic state of Butrinti Lake, demonstrating the influence of climate change in driving this parameter toward eutrophic conditions during 2025**.**

- Analysis of Phytoplankton Community composition based on photosynthetic pigments-Example from Lake Ohrid, Albania

CHEMTAX software software algorithm (Mackey *et al*. 1996; Wright *et al*. 1996) estimates the relative contribution of major algal groups to total chlorophyll *a* (Chl-*a*) by calculating specific pigment-to-Chl-*a* ratios. This approach enables the quantification of dominant phytoplankton taxa and supports the assessment of seasonal and interannual variations in community composition (Fig. 4).

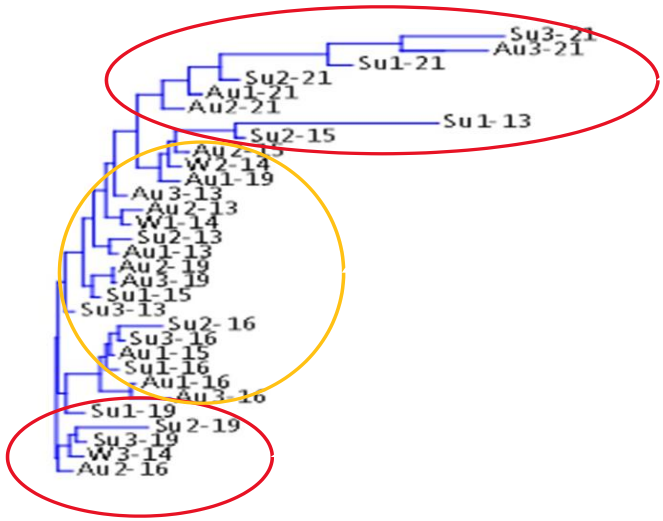


Fig. 4. Dendrogram of similarity among phytoplankton composition (based on pigment ratios chlab/chla, chlc/chla, carotenoids /chla) at three stations (S1-S3) during three seasons (W-winter; Su-Summer; Au-Autumn) at Lake OHRID through the years 2016-2021. Dendrogram was build using Past 4.03. Abbreviation example explanation: Su3-21 stands for Summer-Station3-Year 2021.

The phytoplankton composition observed in 2021 stands distinct from that of all other study years (Fig. 4). The summer–autumn assemblages of 2013, 2014, 2015, and 2019 form a coherent cluster, while 2016 and 2021 each constitute separate groups.

Table 4. An example of Chemotaxonomy based evaluation of the main taxa at Ohrid Lake, based on simplified CHEMTAX approach, applied at samples from 3 stations.

	June 2016 chlb/chla	June 2016 chlc/chla	July 2016 chlb/chla	July 2016 chlc/chla	September 2016 chlb/chla	Sepember 2016 chlc/chla	June 2019 chlb/chla	June 2019 chlc/chla	July 2019 chlb/chla	Jul 2019 chlc/chla	September 2019 chlb/chla	September 2019 chlc/chla
Ohri 1	0.85	1.34	0.97	1.23	0.36	0.75	0.63	0.55	0.86	1.24	0.36	0.75
Ohri 2	1	1.78	0.8	1.05	0.58	0.97	0.88	0.98	0.7	1.02	0.63	0.97
Ohri 3	0.81	1.47	1.18	1.41	0.59	0.97	0.48	0.36	0.82	1.1	0.59	0.97
Ohri1 Taxonomy	*Chlorophyta, Prasinophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Chlorophyta, Prasinophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Prasinophyta, Proclorophyta*	*Dinophyta, Cryptophyta*	*Procolorophyta, Chlorophyta, Euglenophyta, Prasinophyta*	*Dynophyta*	*Chlorophyta, Prasinophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Prasinophyta, Proclorophyta*	*Dinophyta, Cryptophyta*
Ohri 2 Taxonomy	*Proclorophyta, Prasinophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Chlorophyta, Prasinophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Prasinophyta, Proclorophyta*	*Crysophyta, Dinophyta, Cryptophyta*	*Chlorophyta, Prasinophyta, Euglenophyta*	*Dinophyta, Cryptophyta*	*Chlorophyta, Prasinophyta, Proclorophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta*	*Procolorophyta, Chlorophyta, Euglenophyta, Prasinophyta*	*Crysophyta, Dinophyta, Cryptophyta*
Ohri 3 Taxonomy	*Chlorophyta, Prasinophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Proclorophyta, Prasinophyta, Chlorophyta*	*Crysophyta, Dinophyta, Cryptophyta, Haptophyta*	*Prasinophyta, Proclorophyta*	*Crysophyta, Dinophyta, Cryptophyta*	*Proclorophyta, Prasinophyta*	*Cryptophyta, Chrysophyta*	*Chlorophyta, Prasinophyta, Euglenophyta*	*Crysophyta, Dinophyta, Cryptophyta*	*Prasinophyta, Proclorophyta*	*Crysophyta, Dinophyta, Cryptophyta*

This clear differentiation suggests that multivariate analysis of phytoplankton composition can serve as a sensitive indicator for assessing the impact of climate change on aquatic ecosystem dynamics.

- Phytoplankton DNA as a bioindicator of trophic state and exploration of picophytoplankton subspecies diversity

It has been hypothesized that particulate DNA concentrations can provide reliable estimates of living biomass (*Holm-Hansen et al. 1968; Dortch et al. 1983*). Indeed, cellular DNA has been shown to be linearly correlated with cell carbon across various phytoplankton species. Early studies (*Holm-Hansen, 1969*) reported discrepancies between biomass estimates derived from DNA and those obtained from particulate carbon, chlorophyll, or ATP, which were attributed to the presence of substantial fractions of detrital DNA. However, this interpretation was later contested for coastal waters (*Falkowski and Owens, 1982; Dortch et al. 1983*). Subsequent work (*Paul et al., 1985*) demonstrated that particulate DNA correlates closely with particulate organic carbon (POC) and bacterial abundance across diverse aquatic environments. DNA content has thus emerged as a scalable cellular component, co-varying with carbon and nitrogen content in phytoplankton cells. This relationship supports the use of total DNA content as an accurate and independent estimator of cellular carbon biomass in unicellular pelagic phytoplankton (*Veldhuis et al. 2008*). Figures **5** and **6** illustrate applications of phytoplankton DNA as a biomarker for assessing water quality in several Albanian lakes.

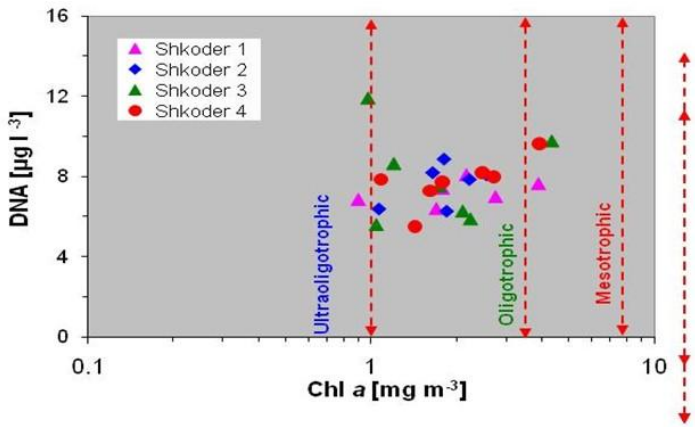


Fig. 5. Evaluation of Trophic State at Lake of Shkodra, Albania based on bioindicators (phytoplankton DNA and chla). Most of the stations have oligotrophic waters.

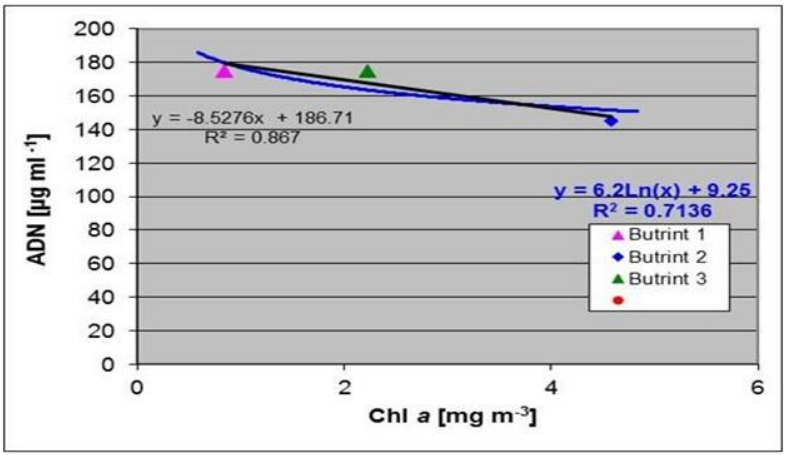


Fig. 6. Correlation between DNA and chla content at Butrinti Lake, Albania. There is a strong correlation R^2=0,867, which proves the possibility to use the phytoplankton DNA as a reliable biomarker for the quality of waters, along with chla.

The quantification of total phytoplankton DNA represents a reliable parameter for assessing phytoplankton biomass, as the cellular DNA content of phytoplankton remains relatively constant under varying environmental conditions. Measurement of this parameter provides an estimate of the number of phytoplankton cells per unit volume of sampled water, independent of their physiological or metabolic activity. Consequently, it reflects the actual biomass of phytoplankton under specific environmental conditions.

In parallel, molecular analyses aimed at exploring subspecies-level diversity were conducted through the amplification of 16S–23S cyanobacterial ribosomal DNA (Fig. 7). This approach has been demonstrated to effectively identify pico-cyanobacteria across diverse aquatic environments worldwide (*Honda et al. 1999; Robertson et al., 2001*).

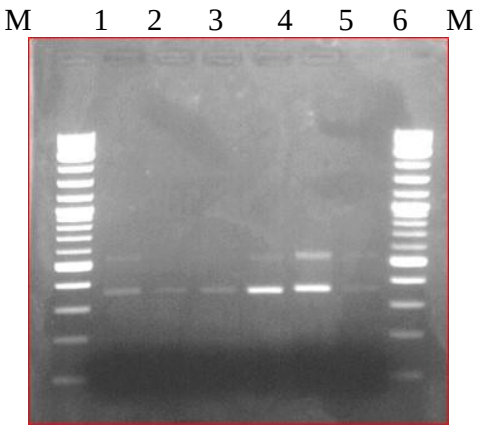


Fig. 7. ITS bf/br 16S-23S ITS rADN amplicon for genus Synechoccocus from Butrinti Lake (Bacu *et al*. 2022a). Primers were designed on the sequence of picophytoplankton 16S-23S rADN according to Lavin *et al*. (2008). Samples of picophytoplankton from different stations have ITS - Internal Transcribed Spacers of two dimensions (350bp and 550bp). From left to right: M-molecular marker 100bp; 1-6 are sampling stations; M-molecular marker 100bp).

- CARD-FISH and Fluorescence Imaging and Flow-Cytometry for phytoplankton biomass evaluation- Case of Lake Butrinti, Albania.

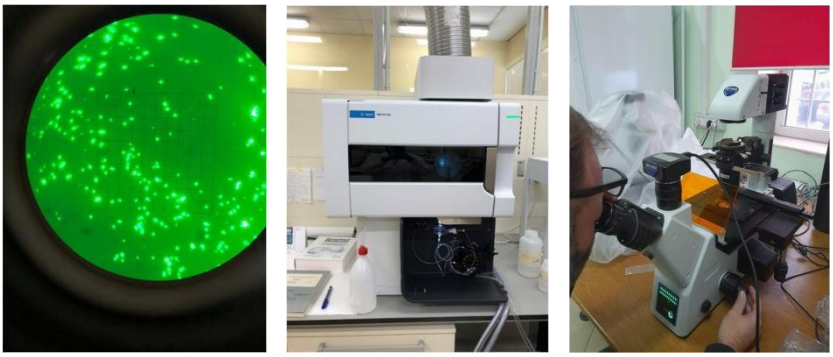


Fig. 8. DAPI staining, Flow-cytometry and fluorescence microscopy were employed to analyze the phytoplankton biomass at Lake Butrinti, Albania in the frame of Bilateral Protocol Italy -Albania 2021-2022. Fram left to right: Image of DAPI staining of filters where phytoplankton was trapped; Flow-cytometer apparatus at Water Research Institute, CNR, Rome, Italy (partner institution); Fluorescence microscope at the Laboratory of Molecular Biotechnology, Department of Biotechnology, University of Tirana, Albania.

A clear dominance of Proteobacteria was observed across all sequenced samples from Lake Butrinti, accounting for approximately 39.4% of the total microbial community. The composition differed notably between water and sediment samples. In water and superficial sediment layers, Bacteroidetes exhibited comparable dominance, contributing up to 30% of the community in some samples. The deep-water microbial community showed greater similarity to that of shallow sediment samples, indicating potential sediment–water microbial exchanges. In surface water samples, Cyanobacteria represented the third most abundant phylum (~10%), whereas their relative abundance dropped below 1% in deeper sediment samples.

4. CONCLUSIONS

Each of the ISO-based methodologies described above provides valuable information on distinct categories of water quality indicators currently applied in Albania. However, a truly integrated assessment framework—combining biotechnological tools such as CARD-FISH**,** fluorescence microscopy, and flow cytometry for biomass evaluation, with factorial analysis of bacterial DNA–environment interactions, phytoplankton genetic diversity**,** and chemotaxonomic profiling—offers a more comprehensive understanding of aquatic ecosystems. Such an approach enables the spatiotemporal analysis of phytoplankton biomass and composition and their correlation with physico-chemical parameters, thus supporting the prediction and interpretation of microbial community dynamics under changing climatic conditions. Furthermore, the incorporation of advanced bioanalytical methodologies (e.g., cell biosensors and chemotaxonomic techniques) facilitates the evaluation of cytotoxic effects on microbiota and associated aquatic food webs caused by pollutants of various origins, including toxic algal blooms and agricultural or industrial discharges. These approaches substantially enhance the resolution and interpretive power of traditional monitoring methods based solely on physico-chemical parameters.

Data accessibility: There are no databases to be made public.

Declaration of AI use:

There has been no use of AI when writing the actual paper.

Author's contribution:

AB - conceptualisation, writing-original, editing draft, evaluation**; XhO**-writing original**; RS-**writing original**; KP-**writing original

Authors consent:

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration: The authors declare there are no conflicts of interest.

Funding: The authors declare no funding was received when writing this paper.

REFERENCES:

Bacu A, Suli K, Zaho I. 2022a. Diversity of picocyanobacteria community and quality of waters at anoxic Lake Butrinti. 4rth International Conference on Applied Biotechnology, Ioannina, Greece. Book of Proceedings, 21-28, ISBN: 978-960-233-276-4.

Bacu A, Zaho I. 2022b. Determining environmental factors that may Influence cyanobacteria bloom formation in Lake Butrinti, Albania. *Journal of Environmental Protection and Ecology*, 23(1): 1–8 (2022). https://scibulcom.net/en/journal/1311-5065/issue/2022-23-1/.

Carlson RE. 1977. A Trophic State Index for Lakes. *Limnology and Oceanography*, **22 (2):** 361–369. doi:10.4319/lo.1977.22.2.0361. ISSN 0024-3590.

Dortch Q, Roberts T, Clayton J, Ahmed S. 1983. RNA/DNA ratios and DNA concentrations as indicators of growth rate and biomass in planktonic marine organisms. *Marine Ecology Progress Series*, **13(1):** 61–71. http://www.jstor.org/stable/24815843.

Falkowski PG, Owens TG. 1982. A technique for estimating phytoplankton division rates by using a DNA-binding fluorescent dye1. *Limnology and Oceanography*, **27 (4):** 776-782. https://doi.org/10.4319/lo.1982.27.4.0776.

Holm-Hansen O, William H, Sutcliffe J, Sharp J. 1968. Measurement of deoxyribonucleic acid in the ocean and its ecological significance.

Limnology and Oceanography, **13 (3):** 507-514 doi: 10.4319/lo.1968.13.3.0507.

Holm-Hansen O. 1969. Determination of microbial biomassin ocean profiles. *Limnology and Oceanograph*y, **14(5):** 740–747. http://www.jstor.org/stable/2834068.

Honda D, Yokota A, Sugiyama J. 1999. Detection of Seven Major Evolutionary Lineages in Cyanobacteria Based on the 16S rRNA Gene Sequence Analysis with New Sequences of Five Marine Synechococcus Strains. *Journal of Molecular Evolution,* **48 (6):**723-39. doi: 10.1007/pl00006517.

ISO 17025- Water quality based on physical-chemical indicators - Environmental Chemistry & Water and Wastewater Treatment Laboratory.

ISO 16649-3:2015- Microbiology of food chain- Horizontal method for the enumeration of beta-glucuronidase-positive *Escherichia coli*- Part 3: Detection and most probable number technique using 5-bromo-4-chloro-3-indolyl-beta-D-glucuronide.

ISO 11348-3:2007- Water quality- Determination of the inhibitory effect of water samples on the light emission of *Vibrio fischeri* (Luminescent bacteria test) - Part 3: Method using freeze-dried bacteria.

Lavin P, Gomez P, Gonzále B, Ulloa O. 2008. Diversity of the marine picocyanobacteria Prochlorococcus and Synechococcus assessed by terminal restriction fragment length polymorphisms of 16S-23S rRNA internal transcribed spacer sequences. *Revista Chilena de Historia Natural.* 81. 515-531. ISSN 0716-078X. http://dx.doi.org/10.4067/S0716-078X2008000400006.

Mackey MD, Mackey D J, Higgins HW, Wright SW. 1966. CHEMTAX- a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. *Marine Ecology Progress* Series, **144(1)**: 265-283. DOI:10.3354/meps144265.

Paul J, Jeffrey W, De Flaun M. 1985. Particulate DNA in subtropical oceanic and estuarine planktonic environments. *Marine Biology*, 90, 95-101. https://doi.org/10.1007/BF00428219.

Robertson BR, Tezuka N, Watanabe M. 2001. Phylogenetic analyses of Synechococcus strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. *International Journal Of Systematic*

And Evolutionary Microbiology, **51 (3)**: 861-871. doi: 10.1099/00207713-51-3-861.

Veldhuis M, Colijn F, Admiraal W. 2008. Phosphate Utilization in *Phaeocystis pouchetii* (Haptophyceae). *Marine Ecology*, **12 (1),** 53 - 62.

Wright S, Thomas D, Marchant H, Higgins H, Mackey M, Mackey D. 1966. Analysis of phytoplankton of the Australian sector of the Southern Ocean: comparisons of microscopy and size frequency data with interpretations of pigment HPLC data using the 'CHEMTAX' matrix factorisation program. *Marine Ecology Progress Series*, **144**: 285-298.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/
which permits unrestricted use, provided the original author and source are credited.

GENETIC POLYMORPHISM OF THE CALPASTATIN (CAST) GENE IN ALBANIAN AND IMPORTED SHEEP BREEDS

Anila HODA
Academy of Sciences of Albania
Sulltane AJÇE
Department of Biochemistry and Agrifood, Faculty of Agriculture, University Fan S. Noli, Korça, Albania
Ilia MIKEREZI
Department of Biology, Faculty of Natural Sciences, University of Tirana, Albania
Academy of Sciences of Albania
Xhelil KOLECI
Department of Veterinary Public Health, Faculty of Veterinary Medicine, University of Agriculture, Tirana, Albania
Nazim GRUDA
Department of Horticultural Sciences, University of Bonn, Germany
Academy of Sciences of Albania
Author for correspondence: anila.hoda@akad.gov.al
AH, 0000-0003-0906-2550; SA, 0009-0009-5737-4461;
IM, 0000-0002-5184-513X;
XhK, 0000-0001-8945-7776; NG, 0000-0002-2663-1594

https://doi.org/10.67537/ajnts3001009

ABSTRACT

Sheep breeding plays a central role in the livestock sector of Albania, contributing significantly to rural livelihoods through the production of meat, milk, and traditional dairy products. However, local sheep breeds often exhibit modest productivity compared with specialized foreign breeds. Advances in molecular genetics have enabled the identification of candidate genes associated with economically important traits, offering opportunities for marker-assisted selection (MAS) in breeding programs. Among these, the calpastatin (CAST) gene is a key regulator of muscle growth and postmortem meat tenderness through its inhibitory effect on calpain proteases. This study investigated the polymorphism of the CAST gene in 112 sheep from two autochthonous Albanian

breeds (Shkodrane and Lara e Polisit) and two imported breeds (Awassi and Île de France) reared at the Centre of Agricultural Technology Transfer in Korça, using PCR–RFLP with the MspI restriction enzyme. A 622 bp fragment encompassing exon 3 was successfully amplified and digested, revealing two alleles (M and N) and two genotypes (MM and MN), while the NN genotype was absent across all breeds. Allele frequencies showed the predominance of M (ranging from 0.865 in Lara e Polisit to 0.981 in Awassi), with breed-specific variations in heterozygosity and fixation indices. The higher genetic diversity detected in the Lara e Polisit breed than in imported populations emphasizes the adaptive potential and conservation value of local Albanian sheep breeds. These findings confirm the polymorphic nature of the CAST gene in Albanian sheep and underline its potential for use in MAS programs aimed at improving growth rate, carcass quality, and meat tenderness, while simultaneously supporting the sustainable conservation of indigenous genetic resources and enhancing competitiveness within the framework of European Union (EU) integration.

Keywords: polymorphism, PCR–RFLP, Albanian sheep breeds, genetic diversity, meat quality, marker-assisted selection (MAS), conservation

1. INTRODUCTION

Sheep breeding plays a vital role in global livestock production, providing meat, milk, and wool as major sources of food and income, particularly in the rural economies of the Western Balkans, especially in Albania. Local breeds in Albania are well adapted to mountainous and semi-arid environments (Hoda *et al*. 2011), however, they often display modest productivity levels compared to specialized foreign breeds. Among the economically important traits in sheep, growth performance and carcass quality are of significant interest to farmers and consumers. These traits are quantitative in nature, influenced by both environmental and genetic factors, and often exhibit low to moderate heritability, which limits the effectiveness of traditional selection methods (Gregula-Kania, 2012). The advent of molecular genetics has enabled the identification of candidate genes and DNA markers associated with production traits, providing opportunities for marker-assisted selection (MAS) to enhance genetic progress in breeding programs.

One of the most widely studied candidate genes in farm animals is calpastatin (CAST), located on chromosome 5 in sheep. CAST encodes calpastatin, an endogenous inhibitor of calcium-dependent proteases, calpains, which regulate muscle protein turnover and play a key role in muscle development and post-mortem meat tenderization (Goll *et al*. 2003;

Yilmaz *et al*. 2024). Increased calpastatin activity reduces protein degradation, which may enhance muscle growth and improve meat quality traits, such as tenderness and carcass composition (Casas *et al*. 2006; Schenkel *et al*. 2006).

Genetic polymorphisms in the CAST gene have been widely documented in sheep populations across different geographical regions. In Turkey (Yilmaz *et al*. 2014; Yilmaz *et al*. 2024), such polymorphisms have been reported in several native breeds, while similar studies in Iran (Azari *et al*. 2012), and Pakistan (Suleman *et al*. 2012), have confirmed variation in local sheep populations. In Europe, genetic diversity at the CAST locus has been identified in breeds such as Polish Merino, Berrichon du Cher, Blackheaded Mutton, and Île de France sheep (Szkudlarek-Kowalczyk *et al*. 2011). Several association studies have shown significant relationships between CAST genotypes and growth or muscling traits (Palmer *et al*. 1999; Byun *et al*. 2008). These findings highlight the global significance of CAST as a candidate gene for small ruminant genetic improvement programs, particularly in relation to growth and meat quality traits.

However, research on genetic diversity and CAST gene polymorphisms in Albanian sheep breeds is absent, despite their socio-economic and cultural importance. Expanding such investigations could provide essential insights for the sustainable genetic improvement and conservation of indigenous Albanian sheep populations, while simultaneously strengthening the competitiveness of the national livestock sector within the framework of European Union (EU) integration.

The study aimed to investigate polymorphisms in the calpastatin (CAST) gene, a potential candidate for future genomic selection, in two autochthonous (Shkodrane and Lara e Polisit) and two imported sheep breeds (Awassi and Île de France) populations raised in Centre of Agricultural Technology Transfer in Korça, using the PCR-RFLP method. Incorporating molecular genetics with performance traits, such as growth and meat quality, can enhance breeding programs for these sheep breeds, while genotyping with CAST markers may allow early classification of carcasses by meat quality before slaughter (Yilmaz *et al.* 2014).

2. MATERIALS AND METHODS

Animals and Sampling

The present study involved two autochthonous sheep breeds (Shkodrane and Lara e Polisit) and two imported breeds (Awassi and Île

de France), all reared at the Centre of Agricultural Technology Transfer in Korça, Albania. In total, 112 blood samples (27–29 individuals per breed) were collected from healthy, unrelated animals of different age categories in the study. For each animal, approximately 5–10 mL of whole blood was obtained via jugular venipuncture and transported on ice to the laboratory for analysis. Genomic DNA was subsequently extracted using the salt-out method described by Gaaib *et al.* (2011).

PCR amplification of CAST gene fragment

A 622-bp fragment of the CAST gene containing the MspI polymorphic site was amplified by polymerase chain reaction (PCR) using primers described by Palmer *et al.* (1998). The forward primer was 5′-TGGGGCCCAATGACGCCATCGATG-3′ and the reverse primer was 5′-GTGGAGCAGCACTTCTGATCACC-3′. PCR amplification was performed in a total reaction volume of 25 µL containing approximately 50 ng of genomic DNA, 1× PCR buffer, 2.0 mM $MgCl_2$, 200 µM of each dNTP, 0.4 µM of each primer, and 1 U of Taq DNA polymerase. The PCR cycle was performed as follows: an incubation step at 95°C for 12 min, followed by 35 cycles of 95°C for 45 s, annealing at 60°C for 45 s and an extraction step at 72°C for 1 min, and 10 min at 72°C as the final extraction.

Genotyping and polymorphism detection

The amplified 622 bp fragment of the CAST gene was genotyped using the restriction endonuclease MspI. For the digestion reaction, a total volume of 50 µL was prepared, consisting of 10 µL of PCR product, 5 µL of 10× One buffer, 0.5 µL of BSA (100×), 0.3 µL of MspI enzyme, and ultrapure water to reach the final volume. The mixture was incubated at 37°C for 1 hour, followed by enzyme inactivation at 65°C for 20 minutes. The digested fragments were separated by electrophoresis on a 3% agarose gel and visualized under ultraviolet (UV) light after ethidium bromide staining. A 50 bp DNA ladder was used as a molecular size marker to determine the fragment sizes generated by PCR-RFLP analysis.

Statistical analysis

Allelic and genotypic frequencies of the CAST gene were estimated using GenAlEx software (Peakall and Smouse, 2012). The conformity of the populations to Hardy–Weinberg equilibrium (HWE) was tested with the Genepop program (Raymond and Rousset, 1995). To evaluate within-

and between-population genetic variation, genetic diversity parameters, and population differentiation indices, we calculated using FSTAT version 2.9.4 (Goudet, 1995). In addition, Nei's gene diversity (Hs), which represents the expected heterozygosity, was computed for each breed to assess its genetic variability.

3. RESULTS

Genotypic structure and allelic frequencies of the CAST Gene

Polymorphism analysis of the CAST gene was performed using the MspI restriction enzyme. The PCR-RFLP results confirmed the presence of genetic variation in all four sheep breeds (Shkodrane, Lara e Polisit, Île de France, and Awassi). Two genotypes were identified following digestion with MspI: MM (336 and 286 bp fragments) and MN (622, 336, and 286 bp fragments). The NN genotype was not detected in any of the examined populations. This locus is therefore characterized by the presence two alleles, M and N. The allelic frequencies of the CAST gene across the breeds are presented in Table 1, showing that the M allele had the highest frequency in all populations, ranging from 0.865 in Lara e Polisit to 0.981 in Awassi sheep, while the N allele was present at relatively low frequencies.

Table 1. Allelic frequencies of CAST gene for four sheep breeds

Locus	Allele	Shkodrane	Lara e Polisit	in Île de France	Awassi
CAST	M	0.889	0.865	0.964	0.981
	N	0.111	0.135	0.036	0.019

Genetic diversity at the CAST locus in four sheep breeds

The genetic diversity of the CAST locus was evaluated in all four sheep breeds. The results of the population genetic parameters are shown in Table 2. The number of effective alleles (Ne) ranged from 1.038 in Awassi to 1.304 in Lara e Polisit, indicating differences in the level of allelic variation among the breeds. The Shannon's information index (I) showed the highest value in Lara e Polisit (0.395), suggesting relatively greater genetic variability in this population than in other breeds, whereas the lowest value was recorded in Awassi sheep (0.092), reflecting a more homogeneous genetic structure.

The observed heterozygosity (Ho) was highest in Lara e Polisit (0.269), followed by Shkodrane (0.148), whereas markedly lower values were found in Île de France (0.071) and Awassi (0.037). In contrast, the expected heterozygosity (He) was generally higher than Ho in most breeds, indicating a deficiency of heterozygotes. The unbiased expected heterozygosity (uHe) confirmed these trends, with values ranging between 0.037 in Awassi and 0.238 in Lara e Polisit cattle. The fixation index (F) varied from –0.156 in Lara e Polisit, indicating an excess of heterozygotes, possibly due to crossbreeding, to 0.250 in Shkodrane, suggesting a higher degree of inbreeding.

These findings highlight considerable inter-population differences in genetic variability at the CAST loci. The higher diversity observed in Lara e Polisit sheep may be linked to their mixed breeding history and adaptability, whereas the low heterozygosity and near fixation of the M allele in imported breeds, such as Awassi and Île de France, reflect a more uniform genetic base.

Table 2. Genetic diversity of CAST gene in four sheep breeds

Breed	**N**	**Ne**	**I**	**Ho**	**He**	**uHe**	**F**
Shkodrane	27	1.246	0.349	0.148	0.198	0.201	0.250
Lara e Polisit	26	1.304	0.395	0.269	0.233	0.238	-0.156
Île de France	28	1.074	0.154	0.071	0.069	0.070	-0.037
Awassi	27	1.038	0.092	0.037	0.036	0.037	-0.019

4. DISCUSSIONS

Genetic characterization of livestock populations using molecular markers has become an essential tool in modern animal breeding, enabling the identification of allelic variations associated with economically important traits, such as growth performance, carcass composition, and meat quality. In this study, polymorphisms in the calpastatin (CAST) gene were investigated in two autochthonous sheep breeds of Albania (Shkodrane and Lara e Polisit) alongside two imported breeds (Awassi and Île de France).

Genetic polymorphisms of the calpastatin (CAST) gene have been extensively studied in sheep populations across different regions, consistently confirming the role it plays in the growth and meat quality

traits. Gregula-Kania (2012) analysed synthetic lines derived from Berrichon du Cher × Polish Merino and Suffolk × Polish Merino crosses, identifying alleles b and e as favourable for growth rate and genotype ac as associated with muscle mass, although no statistically significant correlations with productivity traits were established. In Pakistan, Suleman *et al*. (2012) investigated the exon 1C/1D region of the CAST gene using PCR–RFLP with the MspI enzyme and reported that the MM genotype has the highest frequency, with all three genotypes detected and the M allele showing the highest frequency across Thalli, Lohi, and Kajli breeds. Likewise, Yilmaz *et al*. (2014) studied Kıvırcık, Sakız, Karacabey Merino, and Gökçeada sheep from Western Anatolia, finding high frequencies of the M allele and variable NN genotype distribution, which was particularly high in Sakız but absent in Gökçeada.

Further evidence of the significance of CAST was provided by Byun *et al*. (2008), who reported associations between CAST variants and growth and carcass traits in Romney sheep. In Iranian Afshari sheep, Nikmard *et al.* (2012) reported two alleles at the CAST locus, though without significant association to carcass or growth traits, while Azari *et al*. (2012) found all three genotypes in Dalagh sheep, with deviations from Hardy–Weinberg equilibrium suggesting selection or genetic drift. In Zandi sheep, Khederzadeh *et al*. (2016) observed polymorphism at the CAST locus within a population in equilibrium, confirming the marker's potential in indigenous sheep breeding programs.

Research from the Middle East and Balkans further supports these findings. Jawasreh *et al*. (2017) reported significant associations between CAST polymorphisms and growth and carcass traits in Awassi lambs, with MN genotypes performing better than MM. In Russian half-bred Poll Dorset × North Caucasian meat-wool sheep, Nikolayevna *et al.* (2022) identified MM and MN genotypes, with MM most frequent, while animals with MN showed favourable carcass characteristics. Similarly, Surov *et al*. (2023) highlighted that the MN genotype of CAST, in combination with AB/BB genotypes of the GH gene, was associated with improved growth rate, slaughter weight, and muscle-to-bone ratio. In Norduz sheep from Turkey, Yilmaz *et al*. (2024) confirmed polymorphism at the CAST locus in populations in Hardy–Weinberg equilibrium. In European breeds Dimitrova *et al*. (2017) reported a high frequency of the M allele and predominance of the MM genotype in Karnobat Merino sheep, whereas Bozhilova *et al*. (2022) found CAST polymorphism only in the Synthetic Population Bulgarian Milk breed, while Cooper-Red Shumen and

Karakachan sheep were monomorphic, reflecting restricted diversity due to small population sizes and historical breeding practices.

The results of our study also confirm that the CAST gene is polymorphic in Albanian sheep breeds, with the presence of the M and N alleles and variable genotypic distributions across both autochthonous (Shkodrane, Lara e Polisit) and imported (Awassi, Île de France) populations. Importantly, the higher level of genetic diversity observed in the Lara e Polisit breed, compared with the more uniform structure of imported breeds such as Awassi and Île de France, underscores the adaptive potential and conservation value of local Albanian sheep breeds. This diversity provides a genetic reservoir that can be strategically utilized in marker-assisted selection (MAS) programs to improve traits, such as lamb growth rate, carcass quality, and meat tenderness, while maintaining breed integrity.

From a broader perspective, the identification of CAST polymorphisms in Albanian sheep contributes to the understanding of small ruminant genetic resources in the country, where studies remain limited compared with other countries. Integrating molecular marker information into national breeding programs can strengthen Albania's capacity to enhance productivity, promote sustainable livestock farming, and support alignment of the livestock sector with European Union standards. Such approaches not only improve economic returns for farmers but also play a vital role in safeguarding biodiversity and traditional sheep breeds that are central to the region's cultural and agricultural heritage.

ACKNOWLEDGMENTS

This research was supported by the National Agency for Scientific Research and Innovation (NASRI), Albania, under the Technology and Innovation Projects (TIP) framework, as part of the project titled, Integration of genetic and environmental factors to enhance productivity and health in indigenous sheep.

Consent for publication: Not applicable

Data accessibility: All data generated or analyzed during this study are included in this published article.

Conflict of interest declaration: The authors declare that they have no competing interests.

Declaration of AI use: Artificial intelligence (AI) tools (ChatGPT, OpenAI) were used solely to improve the grammar and readability of the manuscript. All scientific content, analyses, interpretations, and conclusions are the original work of the authors.

Acknowledgement. This research was supported by the National Agency for Scientific Research and Innovation (NASRI), Albania, under the Technology and Innovation Projects (TIP) framework, as part of the project titled "Integration of genetic and environmental factors to enhance productivity and health in indigenous sheep"

Authors' contributions: A. H.: conceptualization, writing—original draft; S.A.: data analysis, validation, writing—review and editing. I. M: data analysis, validation, writing—review and editing.; XH. K..: data analysis, validation, writing—review and editing; N. G: data analysis, validation, writing—review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

REFERENCES:

A**zari MA, Dehnavi E, Yousefi S, Shahmohamadi L. 2012**. Polymorphism of calpastatin, calpain and myostatin genes in native Dalagh sheep in Iran. *Slovak Journal of Animal Science*, **45**:1–6. Available online: https://office.sjas-journal.org/index.php/sjas/article/view/264/252.

Bozhilova M, Dimitrova I, Stancheva N, Teneva A. 2022. Molecular analysis of ovine calpastatin (CAST) and myostatin (MSTN) genes in lambs from three Bulgarian sheep breeds. *Tradition and Modernity in Veterinary Medicine*, **7 1 (12):** 31-37. Available online: https://www.cabidigitallibrary.org/doi/pdf/10.5555/20220144433.

Byun SO, Zhou H, Forrest RHJ, Frampton CM, Hickford JGH. 2008. Association of the ovine calpastatin gene with birth weight and growth rate to weaning. *Animal Genetics: Molecular Genetics and Genomics,*

3**9 (5):** 572–573. PMID: 18513277 DOI: 10.1111/j.1365-2052.2008.01745.x

Casas E, White SN, Wheeler TL, Shackelford SD, Koohmaraie M, Riley DG, Chase CC, Johnson DD, Smith TPL. 2006. Effects of calpastatin and micro-calpain markers in beef cattle on tenderness traits. *Journal of animal science*, **84(3)**: 520–525. https://doi.org/10.2527/2006.843520x.

Dimitrova I, Bozhilova M, Iliev M. 2017. Study of some genes associated with meat productivity in Karnobat Merino sheep breed using PCR-RFLP. *Journal of Agriculture and Veterinary Science* (IOSR-JAVS), **10(8):** 61-5. https://www.iosrjournals.org/iosr-javs/papers/Vol10-issue8/Version-3/M1008036165.pdf.

Gaaib J, Nassief A, Al-Assie AH. 2011. Simple salting – out method for genomic DNA extraction from whole blood. *Tikrit Journal of Pure Science*, **16(2):** 1813-662. https://iasj.rdd.edu.iq/journals/uploads/2024/12/30/4d7d87808b42c8d57d4edd7b48a50475.pdf.

Goll DE, Thompson VF, Li H, Wei W, Cong J. 2003. The calpain system. *Physiological reviews*, **83(3):** 731–801. https://doi.org/10.1152/physrev.00029.2002.

Goudet J. 1995. FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. Journal of Heredity, 86 (6):485–486. https://doi.org/10.1093/oxfordjournals.jhered.a111627.

Gregula-Kania M. 2012. Effect of calpastatin gene polymorphism on lamb growth and muscling. *Annals of Animal Science*, **12 (1):** 63 - 72. DOI:10.2478/v10220-012-0005-7.

Hoda A, Hykaj G, Sena L, Delia (Veizaj) E. 2011. Population structure in three Albanian sheep breeds using 36 single nucleotide polymorphisms. A*cta Agriculturae Scandinavica, Section A, Animal Science*, **61 (1):**12–20. https://doi.org/10.1080/09064702.2010.542250.

Jawasreh KI, Jadallah R, Al-Amareen AH, Abdullah AY, Al-Qaisi A, Alrawashdeh IM, Al-Zghoul MBF, Ahamed MKA, Obeidat B. 2017. Association between MspI calpastatin gene polymorphisms, growth performance, and meat characteristics of Awassi sheep. *The Indian Journal of Animal Sciences*, **87(5):** 635-639. https://doi.org/10.56093/ijans.v87i5.70269.

Khederzadeh S, Iranmanesh M, Motamedi-Mojdehi R. 2016. Genetic diversity of myostatin and calpastatin genes in Zandi sheep. *Journal*

of Livestock Science and Technologies, **4:**45–52. https://journals.uk.ac.ir/article_1381_bdccb36073291951327ad4187f884ae0.pdf.

Nikmard M, Molaee V, Eskandarinasab MP, Dinparast Djadid N, Vajhi AR. 2012. Calpastatin polymorphism in Afshari sheep and its possible correlation with growth and carcass traits. *Journal of Applied Animal Research*, **40 (4):**346–350. https://doi.org/10.1080/09712119.2012.692330.

Nikolayevna SL, Olegovna FI, Vadimovich KD, Vladimirovna SA, Ivanovna DI, Igorevna KN, Caro Petrović V. 2022. Association of Growth Hormone Gene Polymorphisms and Calpastatin Gene with Quality of Sheep Meat. *Pakistan Journal of Zoology*, **55(6):** pp 2797-2804, 2023 DOI: https://dx.doi.org/10.17582/journal.pjz/20220109220137.

Palmer BR, Morton JD, Roberts N, Ilian MA, Bickerstaffe R. 1999. Marker-assisted Selection for Meat Quality and the Ovine Calpastatin Gene. In *Proceedings of the New Zealand Society of Animal Production.* pp. 266–268. New Zealand Society of Animal Production. Available online https://www.nzsap.org/system/files/proceedings/1999/ab99080.pdf.

Palmer BR, Roberts N, Hickford JG, Bickerstaffe R. 1998. Rapid communication: PCR-RFLP for MspI and NcoI in the ovine calpastatin gene. *Journal of animal science*, **76(5):** 1499–1500. https://doi.org/10.2527/1998.7651499x.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*, **28 (9):** 2537–2539. https://doi.org/10.1093/bioinformatics/bts460.

Raymond M, Rousset F. 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, **86 (3):** 248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573.

Schenkel FS, Miller SP, Jiang Z, Mandell IB, Ye X, Li H, Wilton JW. 2006. Association of a single nucleotide polymorphism in the calpastatin gene with carcass and meat quality traits of beef cattle. *Journal of animal science*, **84(2):** 291–299. https://doi.org/10.2527/2006.842291x.

Suleman M, Khan SU, Riaz MN, Yousaf M, Shah A, Ishaq R, Ghafoor A. 2012. Calpastatin (CAST) gene polymorphism in Kajli, Lohi and

Thalli sheep breeds. *African Journal of Biotechnology*, **11 (47):**10655–10660. https://doi.org/10.5897/AJB11.2478.

Surov AI, Skorykh LN, Skokova AV, Omarov AA, Fominova IO. 2023. Polymorphisms of Growth Hormone and Calpastatin Genes and Their Association with Sheep Meat Quality. *Russian Agricultural Sciences*, **49 (2): S333–S338 (2023).** https://doi.org/10.3103/S1068367423080190.

Szkudlarek-Kowalczyk M, Wiśniewska E, Mroczkowski S. 2011. Polymorphisms of calpastatin gene in sheep. *Journal of Central European Agriculture*, **12(3):** 425-432 DOI: https://doi.org/10.5513/JCEA01/12.3.934.

Yilmaz O, Sezenler T, Ata N, Yaman Y, Cemal İ, Karaca O. 2014. Polymorphism of the ovine calpastatin gene in some Turkish sheep breeds. *Turkish Journal of Veterinary & Animal Sciences*, **38 (4):**354–357. DOI 10.3906/vet-1401-13.

Yilmaz O, Çak B, Demirel AF. 2024. Investigation of calpastatin (CAST) gene polymorphism in Norduz sheep by PCR-RFLP method. J*ournal of the Hellenic Veterinary Medical Society*, **75 (2):**7451–7456. DOI:10.12681/jhvms.34752.

© 2025 The Authors.

Published by the Albanian Academy of Sciences under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/ which permits unrestricted use, provided the original author and source are credited.

C O N T E N T S

No	*Author*	*Research Papers*	*Page*
1	**Klaudio Peqini**	Elastic forces explain asymmetry but not the frequency patterns of dipolar field reversals	3
2	**Besian Rama** **Edmond Dushi** **Damiano Koxhaj**	Near real-time moment tensor processing and validation in the Adria–Albanian orogenic contact: an operational framework on the Albanian broadband seismic network	25
3	**Edmond Dushi** **Besian Rama** **Damiano Koxhaj**	Coseismic stress in the 2019 Durrës seismic sequence: Tensor results and early aftershock adjustment	51
4	**Kleona Binjaku** **Elinda Kajo Meçe** **Cecilia Pasquale** **Simona Sacone**	From mono-class to multi-class physics-guided traffic state prediction: A comparative study of Gaussian processes and LSTM networks	73
5	**Ina Papadhopulli** **Martin Çaro**	LEARNPLC: A JAVAFX-based educational platform for ladder logic programming with real-time circuit simulation	99
6	**Alba Haveriku** **Elinda Kajo Meçe**	Albanian sentiment analysis: Exploring machine and deep learning techniques	123
7	**Xhensila Omeri** **Ariola Bacu**	Phytoplankton monitoring and genetic analysis of Lake Butrinti: Advanced methods and environmental implications	133
8	**Ariola Bacu** **Xhensila Omeri** **Rigerta Sadikaj** **Klementina Puto**	An integrated approach based on the use of monitoring and research protocols for water quality assessment in Albania	161
9	**Anila Hoda** **Sulltane Ajçe** **Ilia Mikerezi** **Xhelil Koleci** **Nazim Gruda**	Genetic polymorphism of the Calpastatin (*CAST*) gene in Albanian and imported sheep breeds	177